Spatial and environmental influences on selection in a clock gene coding trinucleotide repeat in Canada lynx (Lynx canadensis)

Melanie Prentice[1], Jeff Bowman[2], Dennis Murray[1], Kamal Khidas[3], and Paul Wilson[1,1]

[1]Trent University
[2]Ontario Ministry of Natural Resources and Forestry
[3]Canadian Museum of Nature

August 23, 2020

Abstract

Clock genes exhibit substantial control over gene expression and ultimately life-histories using external cues such as photoperiod, and are thus likely to be critical for adaptation to shifting seasonal conditions and novel environments as species redistribute their ranges under climate change. Coding trinucleotide repeats (cTNRs) are found within several clock genes, and may be interesting targets of selection due to their containment within exonic regions and elevated mutation rates. Here, we conduct inter-specific characterization of the NR1D1 cTNR between Canada lynx and bobcat, and intra-specific spatial and environmental association analyses of neutral microsatellites and our functional cTNR marker, to investigate the role of selection on this locus in Canada lynx. We report signatures of divergent selection between lynx and bobcat, with the potential for hybrid-mediated gene flow in the area of range overlap. We also provide evidence that this locus is under selection across Canada lynx in eastern Canada, with both spatial and environmental variables significantly contributing to the explained variation, after controlling for neutral population structure. These results suggest that cTNRs may play an important role in the generation of functional diversity within some mammal species, and allow for contemporary rates of adaptation in wild populations in response to environmental change. We encourage continued investment into the study of cTNR markers to better understand their broader relevance to the evolution and adaptation of mammals.

INTRODUCTION

In the near future, global climate change will surpass habitat destruction as the leading threat to biodiversity (Leadley et al., 2010). Given such changes, the ability of species to persist will rely on either redistribution, evolution towards new behavioural and physiological optima, or both. While a number of range redistributions have already been observed (e.g., Laliberte & Ripple, 2004; Parmesan, 2006), the few species that are capable of such large-scale dispersal to track changing environments are unlikely to redistribute to landscapes containing their complete suite of optimal habitat conditions. For example, where species redistribute to maintain equilibrium with one habitat feature (e.g., prey abundance or distribution, temperature), they might simultaneously become discordant with other critical features (e.g., novel predators, photoperiod). Thus, it is likely that for most species, some extent of microevolution will be necessary to allow for their persistence in a changing climate (Visser, 2008).

With the advent of novel molecular technologies and developments in bioinformatics, acquisition and analyses of large-scale data has made the investigation of genomic elements of adaptation possible for both model and non-model species (Harrison, Pavlova, Telonis-Scott, & Sunnucks, 2014). In wild populations, adaptive genomic change will be dependent on the level of standing genetic variation or the rate of *de novo* mutations in the existing population (Carlson, Cunningham, & Westley, 2014). While standing genetic variation is considered most important for contemporary adaptation as it is readily available for selection (Barrett &

1

Schluter, 2008), genetic variation in ecologically important traits may be absent in widely dispersed species, making them targets for extirpation. Specifically, panmictic species experiencing high rates of dispersal and gene flow may be limited in the extent to which adaptive genetic mutations can accumulate and contribute to adaptation. However, while gene flow is predicted to preclude the local adaptation of wide-ranging species (Lenormand, 2002), recent empirical evidence provides support for adaptive genetic divergence in the face of ongoing gene flow (e.g., André et al., 2011; Feder, Egan, & Nosil, 2012; Hemmer-Hansen, Nielsen, Frydenberg, & Loeschcke, 2007; Nielsen et al., 2009). This is often attributed to adaptations along ecological clines, including clines in salinity (Defaveri, Johnsson, & Merilä, 2013), temperature (Aitken, Yeaman, Holliday, Wang, & Curtis-McLane, 2008), or altitude, (Keller, Taverna, & Seehausen, 2011). Characterizing this diversity across species distributions can allow us to better understand the "adaptive potential" of species, and contribute to conservation plans where needed (Keller et al., 2011; Meröndun, Murray, & Shafer, 2019).

In lieu of high standing genetic variability, some species will rely largely on *de novo* mutations for adaptation, a mechanism that may not be effective over short time-frames. Specifically, rates of mutation are likely to be too low (Fan & Chu, 2007), most new mutations will be either neutral or deleterious rather than providing an adaptive benefit (Eyre-Walker & Keightley, 2007), and in small or declining populations, deleterious mutations can accumulate (Willi, Griffin, & Van Buskirk, 2012). Further, the fate of new mutations depends on generation time and population size, where species with longer generation times and smaller population sizes have fewer mutational opportunities (Carlson et al., 2014). In cases where beneficial mutations do arise, they will have to persist until they reach high frequencies, during which time the chances of losing beneficial alleles to genetic drift or demographic stochasticity can be high (Whitlock, 2000).

Considering the importance of rapid adaptation, coding trinucleotide repeats (cTNRs) may be important targets of selection due to their containment within exonic regions of the genome and elevated mutation rates in comparison to more conventionally studied markers (ranging from 10^{-6} to 10^{-2} for microsatellites versus 10^{-9} for SNPs; Fan & Chu, 2007). These repeats are found within genes that code for biologically relevant functions, and may facilitate the fine-tuning of gene function and expression patterns (Gemayel, Vinces, Legendre, & Verstrepen, 2010; Gemayel, Cho, Boeynaems, & Verstrepen, 2012; Haerty & Golding, 2010). Most interestingly, however, is the high mutability of cTNRs, which has been attributed to their inherent instability and may allow for higher rates of evolution in response to environmental stress (Gemayel et al., 2010, 2012). For example, the presence of variable cTNRs in genes coding for body morphology has implicated these repeats as the driver of the rapidly evolved canid skeleton in an evolutionarily short time span of fifty years (Laidlaw et al., 2007). This provides compelling evidence for the role of cTNRs in the rapid morphological adaptation of large mammals.

The candidate gene approach to identifying loci subjected to natural selection consists of the selection of previously characterized genes that are hypothesized to play a role in local adaptation, or genes that may be biologically influenced by environmental heterogeneity (Rellstab, Gugerli, Eckert, Hancock, & Holderegger, 2015). This type of approach is especially appealing for non-model species with uncharacterized genomes, as genomic information from closely related model species (e.g., homologous genes) can be used to select genes of known function and design primers for applications in non-model species of interest (Primmer, Papakostas, Leder, Davis, & Ragan, 2013). Further, the candidate gene approach is beneficial for traits that are not phenotypically apparent but suspected to be of adaptive significance for adaptation to climate change (Keller et al., 2011). The environmental heterogeneity associated with latitude implies that adaptation should involve a range of traits in species spanning latitudinal gradients. For example, adaptation to varying temperatures and photoperiods, which are often used to time life-history events (e.g., breeding and migration), should be of particular importance to many species (O'Malley, Camara, & Banks, 2007). Thus, adaptation of these traits will be critical for species persistence in changing environments.

Although species can track shifts in temperature via redistribution, exposure to novel photoperiods will likely result in challenges for species relying on photoperiodic cues, as photoperiod-dependent behaviours may not change at the same rate as the change in photoperiod that occurs during a range shift, resulting in photoperiodic mismatches (Milligan, Holt, & Lloyd, 2009). As a result, it is possible that the rate of evolution in

2

traits controlling photoperiod-dependent behaviours will be more important than traits involved in temperature sensitivity (Bradshaw & Holzapfel, 2006, 2008), and the ability of species to adapt their behavioural response to new photoperiods, or eliminate photoperiod-dependent behaviours in favour of opportunism, will become critical (Bronson, 2009). In these cases, microevolution of genes responding to photoperiod is likely necessary to ensure species persistence in the face of climate change (Bellard, Bertelsmeier, Leadley, Thullier, & Courchamp, 2012). Species inhabiting seasonal environments often respond to photoperiod cues via circadian clocks (Goldman, 2001), making clock genes good candidates for characterizing the potential genetic responses of species to shifting seasonal conditions and novel environments (Kondratova, Dubrovsky, Antoch, & Kondratov, 2010). cTNRs have been observed in a number of clock genes, and emerging studies have begun to demonstrate the evolutionary and adaptive importance of clock gene cTNRs in a range of species (e.g., Johnsen et al., 2007; Liedvogel, Szulkin, Knowles, Wood, & Sheldon, 2009; O'Malley, Ford, & Hard, 2010; Prentice et al., 2017a).

Our objective was to explore the relationship between space, the environment, and a putatively adaptive cTNR within the candidate clock gene, Nuclear Receptor Subfamily 1, Group D, Member 1 (*NR1D1*), in Canada lynx (*Lynx canadensis*). The *NR1D1* gene is a nuclear receptor that links circadian rhythms to the transcriptional control of metabolic pathways, and functions as a core component of the circadian clock. Its influence on numerous pathways, including memory and learning, behaviour, immune function, metabolism, and mitochondrial content and function (i.e., exercise capacity), suggest that this gene likely plays a role in adaptation to new environments (Everett & Lazar, 2014). Most interestingly, the *NR1D1* gene has demonstrated involvement in circadian regulation of food entrainment (Delezie et al., 2016), and cold tolerance (Gerhart-Hines et al., 2013). The cTNR within the *NR1D1* gene has been previously shown to exhibit signatures of positive selection in Canada lynx (Prentice et al., 2017a). Lynx are a suitable species for investigating selection on clock genes because they show high gene flow at neutral genetic markers across their mainland range (Row et al., 2012), and thus, spatial patterns in NR1D1 are likely to be due to selection rather than neutral processes such as dispersal and gene flow.

Here, we assessed the influence of spatial and environmental patterns on the variation and distribution of *NR1D1* cTNR alleles in Canada lynx, while controlling for the influence of background population structure using a dataset of neutral microsatellite markers. First, we examine inter-specific differences in the distribution and frequencies of *NR1D1* alleles between Canada lynx, and the closely related species, bobcat (*Lynx rufus*) (Johnson & O'Brien, 1997). Although the evolutionary history of Canada lynx and bobcat is unclear, the largely allopatric distributions of these species suggest that they are likely to be differentially adapted to their respective environments. The lynx is largely distributed across the northern boreal, sub-boreal and western montane forests of North America (Quinn & Parker, 1987), occupying a narrower niche than expected if randomly distributed (Peers, Thornton, & Murray, 2012). The bobcat is broadly distributed throughout southern North America in the contiguous USA, and occupies a broader niche than expected (Peers et al., 2012), suggesting that it is a habitat generalist (Anderson & Lovallo, 2003). We hypothesized that the differing habitats used by the two species will result in a divergence in selection pressures experienced by lynx and bobcat, and predicted that we should observe a corresponding divergence between lynx and bobcat alleles at the *NR1D1* cTNR locus [consistent with the findings of Prentice et al., (2017a)]. We also characterized this locus in introgressed lynx and bobcat individuals (Koen, Bowman, Lalor, & Wilson, 2014) to explore the possibility of selection in the lynx-bobcat contact zone. We hypothesized that selection will favour lynx alleles in bobcats inhabiting lynx habitats (i.e., individuals at the north end of the bobcat range). We contrasted genetic differentiation in our candidate cTNR locus and neutral microsatellite datasets to test for differences in pairwise population differentiation at a putatively selected locus versus neutral loci. As differing processes are responsible for influencing neutral versus functional genetic variation (gene flow and genetic drift versus natural selection, respectively), we hypothesized that estimates of differentiation at our putatively selected cTNR locus would be outside of the range reported for neutral loci, representing population pairs that were under either more similar (*NR1D1* F_{ST} < neutral F_{ST}) or different (*NR1D1* F_{ST}> neutral F_{ST}) selection pressures than would be expected based on gene flow alone.. Further, we assessed the association of environmental variables with the observed genetic structure of the *NR1D1* locus, by

assessing relationships between *NR1D1* alleles and environmental and spatial variables, while controlling for background genetic structure with our neutral microsatellite dataset. We hypothesized that environmental heterogeneity is influencing selection at the *NR1D1* locus, and predicted that we would detect associations between environmental variables associated with lynx habitat selection and *NR1D1* alleles, after removing the effects of background neutral genetic structure.

METHODS

Sample collection

The Canada lynx and bobcat analysis was continental, incorporating the entire range of Canada lynx, and the area of range overlap with bobcat at the southern extent of Canada (see Koen et al., 2014). Hide samples (2.5 mm × 2.5 mm) of legally trapped individuals were collected from the North American Fur Auction. Species assignments were initially made by the trappers who harvest the animals, and then confirmed by fur graders at the auction house and our sampling team. In total, we assessed 1,899 Canada lynx samples and 750 bobcat samples, in addition to 2 introgressed lynx and 5 introgressed bobcat samples identified in Koen et al., (2014) (Table 1, Fig. 1). The samples used in this study were a subset of those used in Koen et al., (2014), with the addition of western Canada lynx samples used by Prentice et al., (2017a), and Cape Breton Island lynx samples used by Prentice et al., (2017b).

Neutral and functional genetic datasets

We used an existing data set of 14 neutral microsatellite loci (Fca031, Fca035, Fca077, Fca090, Fca096, Fca441, Fca391, Fca559, Lc106, Lc109, Lc110, Lc111, Lc118) (Koen et al., 2014, Prentice et al., 2017a). We omitted samples missing genotypes from more than 2 loci (Table 1). The candidate gene cTNR amplified in this study is within the *NR1D1* gene. Primer design, optimization and amplification of this marker was conducted in Prentice et al., (2017a). We used the existing dataset of *NR1D1* genotypes and genotyped bobcat individuals and additional lynx from New Brunswick, Cape Breton Island, and Newfoundland according to the same protocol. Sample sizes for analyses of the *NR1D1* locus are slightly smaller due to missing genotypic data for some samples (Table 1).

Inter-specific comparison of the NR1D1 cTNR

Preliminary analyses for signatures of selection conducted on the *NR1D1* locus in Canada lynx and bobcat were conducted by Prentice et al., (2017a). Here, we mapped the distribution and frequencies of *NR1D1* alleles both across and within species, and assessed introgressed lynx-bobcat individuals to determine whether these individuals had lynx, bobcat or mixed genotypes. To facilitate comparison at the *NR1D1* locus, we also assessed allele frequency distributions at neutral microsatellites of lynx and bobcat. All bobcat and introgressed individuals were removed from the remainder of the analyses.

Genetic structure of neutral microsatellites and the NR1D1 cTNR

We first calculated pairwise estimates of Jost's D at both the *NR1D1* locus and our 14 neutral microsatellite loci with the software package diveRsity v1.9.90 (Keenan, McGinnity, Cross, Crozier, & Prodöhl, 2013) implemented in R (R Core Team, 2016). We assigned individuals to populations based on previously reported STRUCTURE results [Newfoundland (Row et al., 2012), Quebec lynx located south of the St. Lawrence River (Koen, Bowman, & Wilson, 2015), Cape Breton Island (Prentice et al., 2017b), and mainland Canada (the remaining distribution of lynx across continental Canada)]. To estimate whether pairwise genetic differentiation at the *NR1D1* locus fell within or outside the expected range of neutrality, we plotted pairwise Jost's D values estimated for our neutral marker dataset with standard error of the mean intervals, and then overlaid each pairwise estimate of Jost's D for the *NR1D1* locus. We used Jost's D here as it uses the effective number of alleles (Kimura & Crow, 1964), which is a more appropriate approach for outlier-based assessments compared to F_{ST} and its analogues, which are dependent on, and can be biased by, within-population heterozygosity (Hedrick, 1999). Further, Jost's D has been shown to perform better at estimating the differentiation in allele frequencies among populations, whereas F_{ST} is preferential for describing the effect of demography on genetic variation (Meirmans & Hedrick, 2011).

4

Environmental associations of the NR1D1 cTNR

To determine the effect of the environment on functional genetic variation at the *NR1D1* cTNR locus across the range of Canada lynx, we conducted multiple constrained redundancy analyses (RDA) testing the relationship between functional genetic variation (*NR1D1* alleles), space, and environmental variables, while controlling for neutral genetic variation (using alleles at our 14 neutral microsatellites). Coordinates for our fur samples were taken as centroids of regionally managed areas that differed between province, territory and state (referred to herein as "sampling unit"). Resolution of sampling units varied by necessity due to different jurisdictional lynx management practices. Lynx from Quebec, Ontario and British Columbia were characterized by centroids of traplines or trapping units called "Unités de gestion des animaux à fourrure" (UGAF) in the case of Quebec. Samples from Alberta were characterized by centroids of fur management units. Samples from Manitoba were largely characterized by centroids of major trapline units, except for samples for which finer spatial resolution (minor trapline units) was available. Similarly, lynx from Alaska were characterized by centroids of minor drainage units when possible, or major drainages when finer scale data was not available. Yukon lynx were characterized by centroids of aggregated traplines provided by the Yukon government. In Atlantic Canada (Newfoundland, Labrador, New Brunswick and Cape Breton Island), samples were characterized by centroids of each of the four locations.

Our selection of environmental variables was based on previously published evidence of lynx occurrence and habitat selection across their range. Within each environmental variable we expect to characterize a range of variation across the large geographic extent of our study that might lead to differential selection in locations where lynx are likely to be present. For example, lynx occupancy modelling suggests that variables including maximum temperature of the warmest month, snow depth and ecoregion are influential in predicting lynx presence/absence (Peers et al., 2012; Peers, Thornton, & Murray, 2013). Further, throughout their eastern distribution, lynx occurrence has been correlated with average annual snowfall, where lynx are unlikely to occur in habitats with <270cm of snow per year (Hoving, Harrison, Krohn, Joseph, & O'Brien, 2005), suggesting that snowfall patterns may drive differential selection for lynx throughout their distribution. We selected 9 variables hypothesized to be important for describing lynx neutral and functional genetic structure including; (1) human influence, (2) snow cover, (3) snow depth, (4) mean temperature of the warmest quarter, (5) mean temperature of the coldest quarter, (6) precipitation of the wettest quarter, (7) precipitation of the driest quarter, and proportions of (8) suitable and (9) unsuitable lynx habitat. To estimate the proportion of suitable and unsuitable lynx habitat, we used a categorical raster layer of land cover (ESA GlobCover Project, 2009), which we reclassified into a binary layer of "lynx habitat suitability" representing land cover types that have been shown to be selected by lynx as suitable habitat [including closed (>40%) needle leaved evergreen forest (>5m), open (15-40%) needle leaved deciduous or evergreen forest (>5m), and closed to open (>15%) mixed broadleaved and needle leaved forest (>5m); supported by lynx selection for coniferous forest, and avoidance of landscapes dominated by deciduous forest; Hoving et al., 2005; Koehler, 1990; Poole, 2003; Walpole, Bowman, Murray, & Wilson, 2012], and unsuitable habitat (all other land cover types).

The spatial resolution of our environmental variables ranged between 300m - 1km, with varying degrees of temporal resolution (detailed descriptions of variables in Table 2). As the effects of neutral differentiation over space (e.g., IBD) were accounted for using neutral microsatellite markers, we used latitude and longitude as predictor variables, to represent unaccounted for environmental variation correlated with space such as large-scale climate gradients and photoperiod.

We characterized each sample with environmental data by extracting point estimates (snow cover, snow depth, and mean temperature of the warmest and coldest quarters), averages (human influence), and proportions (suitable and unsuitable lynx habitat) using ArcMap version 10.4.1. Point estimates were extracted at each sample coordinate. Averages were taken over the corresponding sampling unit containing each sample with the exception of Alaska, for which we were unable to obtain geographic raster layers for the major and minor drainage units and instead extracted averages over larger game management subunits (which contain the major and minor drainages). Proportions were extracted for each sample by estimating the total area classified as either suitable (1) or unsuitable (0) lynx habitat and dividing these estimates by the total area of

5

each sampling unit. For the extraction of both averages and proportion estimates, we resampled our raster's to the layer with the lowest resolution, and snapped these raster's and the sampling unit layers together, to align the cells for consistent extraction of estimates across layers.

We conducted a principal component analysis (PCA) in R (R Core Team, 2016) on our set of environmental variables (excluding latitude and longitude) to condense our dataset into fewer, uncorrelated variables describing environmental variation across the landscape. We retained principal components that explained >80% of the cumulative proportion of variation, and assessed the eigenvectors of each retained component to determine the variables that loaded most strongly. We calculated the standardized scores of each retained principal component for subsequent analyses.

We conducted constrained RDA analyses in the R package vegan (Oksanen et al., 2017), where our dependent variables represented "functional genetic diversity" (*NR1D1* alleles), and our independent variables included descriptors of the "environment" (retained environmental principal components), "space" (latitude and longitude), and "neutral genetic structure" (14 neutral microsatellite loci). First we conducted an RDA on the full model, to estimate the significance of, and amount of variation explained by the model including environment, space and neutral genetic structure. We then conducted 3 partial RDAs to estimate the significance of, and variation explained by: (1) spatial variables, (2) environmental variables and (3) neutral genetic structure, while controlling for the influence of the additional variables (e.g., for neutral genetic structure we controlled for spatial and environmental effects). These partial RDAs control for the influence of either environment, space or neutral genetic structure by first removing their effects from the functional genetic dataset, and then performing an RDA on the residual matrix (according to the procedure outlined in Oksanen 2012). Significance of each model was assessed by permutation tests, where the *NR1D1* matrix was randomly permuted 999 times, and the strength of the relationship between the observed or permuted matrices and the independent variables were compared ($\alpha = 0.05$). It has been shown that even if environment and spatial patterns perfectly explain genetic variation, the total inertia (variance) explained can be substantially less than 1 (Økland 1999). Because of this, we followed the recommendations of Økland (1999), and partitioned the variance based on the total *explainable* variance in our full model. To do so we used the full model to estimate the total explainable variation of our combined datasets, and the conditional models (partial RDAs) to partition the total explainable variation into 4 components: (1) variation attributed purely to environment, (2) variation attributed purely to space, (3) variation attributed purely to neutral genetic structure and (4) joint (collinear) variation attributed to environment, space and neutral genetic structure.

As we had a large dataset over a large geographic area which might have obscured environmental gradients of interest, we also reduced our dataset into two halves ("eastern lynx" including individuals sampled within or east of Manitoba (N= 1,466), and "western lynx" including individuals sampled within or west of Alberta/ Yukon (N= 423); Fig. 1A), and repeated the analysis on each subset independently to reduce unexplained environmental variation in the data when analyzed from a continental scale. We were specifically interested in the eastern lynx dataset, where preliminary analyses showed indications of selection in peripheral and insular populations of lynx (Koen et al., 2015, Prentice et al. 2017b). Further, Row et al. (2014), reported a correlation between neutral genetic variability and a winter climate gradient in lynx across the Pacific-North American (PNO) and North Atlantic Oscillation (NAO) climatic systems in eastern Canada, and suggested restricted dispersal in lynx between Manitoba and Quebec supporting the potential for selection in this region.

RESULTS

Inter-specific comparison of the NR1D1 cTNR

We have previously reported a non-overlapping allelic range between Canada lynx and bobcat individuals (excluding introgressed individuals) and signatures of positive selection in Canada lynx using F_{ST} outlier-based analysis (Prentice et al. 2017a). Here, across a much larger-scale screening of lynx and bobcat individuals, we identified the same 7 alleles at the NR1D1 locus with the same non-overlapping patternAll identified bobcat samples had genotypes comprising the smaller 3 *NR1D1* alleles and all identified Canada lynx samples had

genotypes comprising the larger $4NR1D1$ alleles with the exception of 10 putatively introgressed individuals (described below). Gradients in allele frequencies were observed in both lynx and bobcat (Fig., 2). In lynx, allele frequencies were largely similar across the range we surveyed, with the exception of the most southern population sampled (New Brunswick lynx and Quebec lynx sampled south of the St. Lawrence River), which showed a substantial increase in the frequency of the smallest allele observed in lynx (Fig., 2B). Insular populations of lynx (Newfoundland and Cape Breton Island) were fixed for the most common lynx allele (Fig., 2B). Of the 3 alleles identified in bobcats, the intermediate allele was found most prevalently across the range we surveyed whereas the smallest and largest allele were observed most prevalently (although not exclusively) in western and eastern bobcat, respectively (Fig., 2B). In comparison, neutral microsatellite markers were found to have largely overlapping allelic ranges across the two sister species; on average, 42.7% of alleles per locus overlapped, across all 14 presumably neutral loci. We note however, that the range of overlap was variable across loci (from 0-83.3%), with 2/14 loci showing no overlap in alleles between lynx and bobcat.

We further identified 10 putative Canada lynx samples with mixed genotypes (i.e., 1 "bobcat" and 1 "lynx" allele; Fig 1B). These samples were all visually identified as Canada lynx, and contained neutral microsatellite profiles and control region mtDNA sequences consistent with Canada lynx. In addition, we evaluated samples from introgressed animals previously identify by Koen et al., (2014). Of the 5 samples they identified as introgressed bobcats, 4 had genotype profiles consistent with bobcats at the $NR1D1$ locus (i.e., they had 2 "bobcat" alleles), and 1 individual failed to amplify. Of the 2 samples identified by Koen et al., (2014) as introgressed lynx, 1 had a genotype profile consistent with lynx at the $NR1D1$ locus (2 "lynx" alleles), and the other had a mixed profile (1 "lynx" and 1 "bobcat" allele).

Genetic structure of neutral microsatellites and the NR1D1 cTNR

Our Jost's D plot comparing neutral microsatellites and the $NR1D1$ locus showed that the $NR1D1$ locus consistently fell outside of the expected range of neutrality estimated with our 14 neutral microsatellites, suggesting that this locus may be under selection in lynx across their range (Fig., 3). Further, estimates of Jost's D at the $NR1D1$ locus for all pairwise comparisons between Cape Breton Island, Newfoundland, and the mainland population of lynx were lower than neutral expectations, suggesting that they are much more similar than expected based on gene flow. Alternatively, all comparisons including lynx sampled south of the St. Lawrence River were higher than neutral expectations, suggesting that they are much more different than expected based on gene flow alone (Fig., 3).

Environmental associations of the NR1D1 cTNR

We retained the first four principal components for our ecological association RDA analyses, which explained ~86% of the variation in our environmental variable dataset (Table 3). According to the eigenvectors, the first principal component was most highly influenced by precipitation variables (precipitation of the wettest quarter, precipitation of the driest quarter and average snow cover; Table 4). Our second component was largely influenced by habitat (proportion of suitable and unsuitable lynx habitat; Table 4). The third component was influenced by both temperature (mean temperature of the warmest quarter) and average snow depth (Table 4). Our fourth component was explained by a combination of precipitation (precipitation of the wettest and driest quarters, and average snow depth) and temperature (mean temperature of the warmest quarter; Table 4).

The partial RDA models describing environmental and spatial effects on $NR1D1$ genetic variation (controlling for neutral genetic structure) were only significant in the eastern Canada lynx data subset (p=0.001 and 0.008 for space and environment, respectively; Table 5). In this data subset, the total explainable variation was 0.2128, of which the effects of purely neutral genetic structure accounted for the largest proportion of variance (63.20% of total explainable variation, p=0.001), followed by the joint (collinear) effects of neutral structure, environmental and spatial data (28.81% of total explainable variation, p=0.001), with purely environmental and spatial data contributing much smaller, but still significant proportions of variation (4.93% and 3.05% total explainable variation in space and environment, respectively; Table 5). In the full dataset, only the full

7

model and partial RDA model describing the effects of neutral genetic structure (controlling for environmental & spatial data) were significant (p=0.001; Table 5). Here, the total proportion of explainable variation was lower (0.1592), and largely comprised of the effects of purely neutral structure (71.73%), and the joint effects of neutral structure and environment/space (26.01%), with environmental/spatial effects contributing small, insignificant amounts to explained variation (0.19% and 2.07% for space and environment, respectively; Table 5). No models in the western data subset were significant suggesting that neither neutral genetic structure nor environmental or spatial variation explain genetic variation at the *NR1D1* locus in this dataset (Table 5).

DISCUSSION

Multiple analytical approaches suggest the influence of selection at the *NR1D1* locus in Canada lynx, a widely dispersing, broadly-distributed mammal. The non-overlapping allelic ranges observed here between Canada lynx and bobcat at this cTNR might be the result of selection favouring different allele length variants in sister species with largely allopatric distributions. The high mutability of cTNR loci support the possibility that convergent alleles have manifested between these closely-related species; however, our failure to characterize large numbers of lynx alleles in bobcats and vice versa given our large sample sizes concentrated around the area of range overlap suggests that selection pressures may have prevented these alleles from persisting in either species and reaching high frequencies. In contrast, we observed high degrees of overlap in most of our neutral microsatellite markers, however, we note that 2/14 of these presumably neutral loci showed no overlap in alleles between lynx and bobcat. Thus, we cannot exclude the possibility that patterns of non-overlapping allelic ranges in both functional and neutral loci could be the result of drift occurring in largely allopatric species. Alternatively, some of our presumably neutral loci may be in linkage with loci under selection. Ultimately, the use of additional analytical assessments of the *NR1D1* cTNR are critical to substantiate the role of selection on this locus.

We observed gradients in allele frequencies at the *NR1D1* locus within each species, suggesting that variations in environmental and/or spatial patterns could be influencing selection intra-specifically. For example, our most southern sampled lynx population (lynx sampled south of the St. Lawrence River in New Brunswick and Quebec), showed a substantial increase in the frequency of the smallest allele observed in lynx. This is concordant with a latitudinal gradient of allele frequencies in lynx, where the smaller lynx-specific alleles (those closest to the range of alleles observed in southern-adapted bobcats) are found more frequently in more southern distributed lynx. Characterization of lynx at their southern range periphery (i.e. throughout the contiguous United States) would help further elucidate this pattern.

All 10 of the samples we identified with both lynx and bobcat alleles (0.004% of sampled individuals) were within or in close proximity to the area of range overlap between lynx and bobcat (note that the individual sampled in Alberta lacked associated spatial information and was thus plotted as the centroid of the province). As these samples were all otherwise characterized as Canada lynx, our finding could result from 1 of 2 mechanisms (or a combination of both): (1) bobcat alleles are being introduced into southern lynx populations via multiple generations of introgression, and/or (2) convergent evolution of *NR1D1* alleles in lynx, such that the deletion of repeats has resulted in the introduction of "bobcat" alleles into lynx populations via mutation without introgression. The maintenance of these convergent "bobcat" alleles in lynx may then have been maintained via selection in individuals at the southern range periphery. While it may be predicted that bobcat alleles should be selectively disadvantageous to lynx throughout their range, they may be retained in peripheral populations if they provide a competitive advantage or adaptive benefit to environmental conditions at the southern range edge.

We found that almost all of the known introgressed lynx and bobcats identified by Koen et al., (2014) had genotype profiles at the *NR1D1* cTNR locus consistent with their previously identified parental species. While some remained close to the area of overlap, all of these samples were trapped outside of the currently recognized range overlap of lynx and bobcat, supporting the hypothesis that selection favours the retention of pure genotypes comprised of alleles belonging to the parental species. The only introgressed sample containing a mixed profile was that of a lynx trapped in Saskatchewan, however, location data for this

8

sample was imprecise, limiting our interpretation.

While the observation of a split lynx-bobcat profile at the *NR1D1* locus in 1 known introgressed individual provides evidence for introgression, the occurrences of split profiles in 10 individuals otherwise characterized as lynx also supports the convergence hypothesis. These alternative explanations could be more explicitly tested by genotyping historical specimens in areas where lynx and bobcat ranges did not overlap, or characterizing species-specific sequence variants in the flanking sequence of the *NR1D1* locus. Additionally, characterization of bobcats throughout their southern distribution may allow for the differentiation of these two competing hypotheses. For example, if lynx alleles are found to persist in southern distributed bobcats, introgression would be less probable.

The results of introgression of lynx and bobcat alleles could have significant implications for conservation. The overall success of hybridization and introgression induced by range expansions is often driven by interactions between biotic and abiotic factors on multiple spatial and temporal scales (Muhlfeld et al., 2014), the outcomes of which can carry both positive and negative effects. In a positive light, hybridization and introgression can increase the adaptive potential of closely-related species through periods of climate change (e.g., Becker et al., 2013). Alternatively, hybridization may have irreversible evolutionary consequences by reducing the fitness of parental species through the erosion of pure genomes and disruption of co-adapted gene complexes that have evolved together over thousands of generations. The extent to which either of these alternatives apply to the conservation of Canada lynx and bobcat remains unknown, but only low levels of hybridization and introgression have been detected between these species to date (Koen et al., 2014). The expansion of bobcats northward, and subsequent hybridization and introgression with lynx, however, may allow for the persistence and continued range expansion of bobcats by introducing more "northern adapted" lynx alleles into the bobcat gene pool at the *NR1D1* locus and other loci of adaptive importance.

Within Canada lynx, several indicators of selection on the *NR1D1* cTNR persisted. Using our Jost's D outlier approach, all pairwise comparisons showed estimates of Jost's D at the *NR1D1* locus that were outside of neutral expectations of gene flow. Specifically, estimates of Jost's D for pairwise comparisons between Cape Breton Island, Newfoundland and mainland Canada were more similar than expected based on neutral expectations of gene flow (Fig., 3), suggesting convergence across these populations. Alternatively, all pairwise comparisons involving lynx sampled south of the St. Lawrence River were higher than neutral expectations, suggesting that this population of lynx is more genetically different than would be expected by gene flow alone (Fig., 3), and supporting divergent selection of this population despite ongoing gene flow.

Although we did not detect any environmental associations in our full dataset, this may have been caused by noise generated from our large geographic scale that may mask subtler environmental associations, which can vary with spatial scale (e.g., Kozakiewicz et al., 2019). We also note that there was considerable variation in the size and amalgamation of sampling units used here (i.e., traplines, drainages etc.), leading to inconsistent sampling error across our study area. Thus, in some areas we may not have been able to detect true environmental associations due to a lack of resolution.

The significant influence of space and environment in our analysis when the effects of neutral genetic variation were removed suggests that selection may be influencing the pattern of genetic variation at the *NR1D1* locus beyond background population structure and gene flow. This pattern was only evident in the eastern half of our study area, and appeared to be especially influenced by *NR1D1* allele frequencies in peripheral areas of the east. We note, however, that although significant, spatial and environmental variables accounted for the smallest proportion of total explainable variation (~8% combined), and neutral genetic structure (accounting for 63.2% explainable variation) is still the most important factor contributing to the variation in *NR1D1* alleles in eastern lynx. This suggests that the biological relevance of the variables we have chosen for our environmental datasets may be limited in respect to the *NR1D1* locus.

Overall, a large proportion of variation in the *NR1D1* locus remained unexplained using our neutral structure, environmental and spatial datasets, ranging from 74.5% in the western data subset to 84.1% in the full dataset. While most of this is likely to be attributed to the methodology we used (i.e., polynomial distorti-

ons and a lack-of-fit of our data to the response model; Økland 1999) it is also possible that environmental variables unaccounted for in our analyses might influence selection for *NR1D1* alleles in lynx. For example, long-lived mammals occupying higher latitudes are particularly vulnerable to climate change, as they often rely on photoperiod variability to cue proactive responses to changing seasons rather than responding reactively (Bradshaw & Holazpfel, 2008; Bronson, 2009). Photoperiod can strongly influence circadian activity in Canada lynx (Kolbe & Squires, 2007) and the closely related Iberian lynx (*Lynx pardinus* , Beltrán & Delibes, 1994). As clock genes are largely regulated by external cues including photoperiod, the evolution of clock gene cTNRs would be necessary to allow for photoperiodic response mechanisms of lynx where photoperiod and temperature regimes become uncoupled under a changing climate. Other spatially structured drivers of selection could include the NAO and PNO climatic systems, which are associated with the spatial genetic structure of Canada lynx at neutral markers (Row et al., 2014), and may also influence selection in unknown ways. Further, Agee (2000) reported significant clines in habitat quality and availability for lynx across their range. Given the significance of space in our eastern lynx dataset, some of these unmeasured, spatially correlated variables may be influencing genetic variation at the *NR1D1* locus.

One of the difficulties in detecting signatures of selection in cTNR loci is that of the multiple emerging software platforms for identifying outlier loci (i.e., F_{ST} or environmental association approaches), most have been developed for large SNP datasets where tens- to hundreds-of-thousands of loci are analyzed simultaneously. Thus, such approaches are likely to lack power for analyses on datasets of a smaller size, or of single loci. Here, we investigate the potential for selection on the *NR1D1* locus using comparisons of genetic differentiation between neutral and functional genetic datasets, and an ordination analysis (RDA). One of the major benefits of ordination analyses is that they do not require assumptions such as HWE or LD to be met, which are generally violated in loci under the influence of selection. Although our analysis doesn't directly rely on assumptions of LD, however, we cannot exclude the possibility of selection acting on a locus in linkage with the *NR1D1* gene, rather than directly on our cTNR marker itself. An additional benefit of our conditional RDA analysis was our ability to control for background neutral population structure when assessing environmental associations between one or few outlier loci.

Unlike analyses that identify selection using differentiation-based approaches (i.e., F_{ST} outliers), a drawback of environmental association analyses arises because evolution does not necessarily follow environmental change perfectly (Merilä & Hendry, 2014). Thus, true environmental associations may not be identified, even if they are responsible for driving the observed changes in functional genetic diversity. Alternatively, our analysis may have excluded important environmental variables that are associated with the distribution of alleles at the *NR1D1* cTNR locus, or been too coarse to detect a relationship. Thus, the use of finer resolution data and/or the investigation of additional environmental variables shown to influence lynx population structure and habitat selection may help explain greater variation in genetic structure at the *NR1D1* and other cTNR loci.

CONCLUSIONS

As global climate change becomes an increasing threat to biodiversity, our understanding of the adaptive potential of species and the integration of this information into management strategies will be critical in preventing species loss. Thus, determining the rates and mechanisms of species responding to environmental change is an important goal for the conservation of species at risk. Our use of a candidate gene approach allowed for the *a priori* selection of a gene with characterized functions, which we hypothesized to be under selection in our species of interest, Canada lynx. We detected multiple signatures of selection in both an intra-specific comparison of closely-related species (Canada lynx and bobcat), and inter-specifically in Canada lynx at the *NR1D1* clock gene cTNR. The possibility for hybrid-mediated gene flow via introgression between lynx and bobcats can erode the pure genomes of the parental species, and can carry positive and/ or negative consequences by, for example, facilitating the northward range expansion of bobcats. The small yet significant influence of both environment and space on this clock gene locus within the eastern distribution of lynx suggests that environmental factors are influencing the distribution of genetic variation at this locus, some of which may be tightly correlated with space (i.e., latitude). More broadly, signatures of selection on

10

the *NR1D1* locus within lynx suggests that cTNR loci may play a role in generating functional diversity within mammals that can be influenced by selection. Ultimately, the high mutation rates of cTNR loci may allow for contemporary rates of evolution in wild populations. While selection on cTNR loci has already been suggested in wild populations of birds (Johnsen et al., 2007; Liedvogel et al., 2009) and fish (e.g., O'Malley et al., 2010), to our knowledge this is the first report that provides evidence for selection on a cTNR locus in a highly vagile wild mammal. Although cTNRs are largely under-represented in the wildlife literature, they may be important targets of study for those aiming to understand the rapid adaptive potential of species under environmental change.

ACKNOWLEDGEMENTS

Literature Cited:

Agee, J. K. (2000). Disturbance ecology of North American boreal forests and associated norther mixed/subalpine forests. In L. F. Ruggiero, K. B. Aubry, S. W. Buskirk, G. M. Koehler, C. J. Krebs, K. S. McKelvey, & J. R. Squires (Eds.), *Ecology and Conservation of Lynx in the United States* (pp. 39-82). Boulder, Colorado: University Press of Colorado.

Aitken, S. N., Yeaman, S., Holliday, J. A., Wang, T, & Curtis-McLane, S. (2008). Adaptation, migration or extirpation: climate change outcomes for tree populations. *Evolutionary Applications* , 1, 95-111. doi: 10.1111/j.1752-4571.2007.00013.x

Anderson, E. M., & Lovallo, M. J. (2003). Bobcat and lynx. In: Feldhamer, G. A., Thompson, B. C., & Chapman, J. A. (Eds.), *Wild Mammals of North America: Biology, Management and Conservation* (pp. 758-786). Baltimore, Maryland: Johns Hopkins University Press.

André, C., Larsson, L. C., Laikre, L., Bekkevold, D., Brigham, J., Carvalho, G. R., . . . Ryman, N. (2011). Detecting population structure in a high gene-flow species, Atlantic herring (*Clupea harengus*): direct, simultaneous evaluation of neutral vs. putative selected loci. *Heredity* , 106, 270-280. doi: 10.1038/hdy.2010.71

Barrett, R. D. H., & Schluter, D. (2008). Adaptation from standing genetic variation. *Trends in Ecology and Evolution* , 23, 38-44. doi: 10.1016/j.tree.2007.09.008

Becker, M., Gruenheit, N., Steel, M., Voelckel, C., Deusch, O., Heenan, P. B., . . . Lockhart, P. J. (2013). Hybridization may facilitate *in situ* survival of endemic species through periods of climate change. *Nature Climate Change* , 3, 1039-1042. doi: 10.1038/NCLIMATE2027

Bellard, C., Bertelsmeier, C., Leadley, P., Thuiller, W., & Courchamp, F. (2012). Impacts of climate change on the future of biodiversity. *Ecology Letters* , 15, 365-377. doi: 10.1111/j.1461-0248.2011.01736.x

Beltrán, J. F., & Delibes, M. (1994). Environmental determinants of circadian activity of free-ranging Iberian lynxes. *Journal of Mammalogy* , 75, 382-393. doi: 10.2307/1382557

Bradshaw, W. E., & Holzapfel, C. M. (2006). Evolutionary response to rapid climate change. *Science* , 312, 1477-1478. doi: 10.1126/science.1127000

Bradshaw, W. E., & Holzapfel, C. M. (2008). Genetic response to rapid climate change: it's seasonal timing that matters. *Molecular Ecology* , 17, 157-166. doi: 10.1111/j.1365-294X.2007.03509.x

Bronson, F. H. (2009). Climate change and seasonal reproduction in mammals. *Philosophical Transactions of the Royal Society B: Biological Sciences* , 364, 3331-3340. doi: 10.1098/rstb.2009.0140

Carlson, S. M., Cunningham, C. J. & Westley, P. A. H. (2014). Evolutionary rescue in a changing world. *Trends in Ecology and Evolution* , 29, 521-530. doi: 10.1016/j.tree.2014.06.005

Defaveri, J., Jonsson, P. R. & Merilä, J. (2013). Heterogeneous genomic differentiation in marine threespine sticklebacks: adaptation along an environmental gradient. *Evolution* , 67, 2530-2546. doi: 10.1111/evo.12097

Delezie, J., Dumont, S., Sandu, C., Reibel, S., Pevet, P. & Challet, E. (2016). Rev-erbα in the brain is essential for circadian food entrainment. *Scientific Reports* , 6, 19386. doi: 10.1038/srep29386

European Space Agency & Université Catholique de Louvain (2010). GlobCover Project, Version 2.3, 2009. . Accessed 27-10-2016.

Everett, L. J., & Lazar, M. A. (2014). Nuclear receptor Rev-erbα: up, down, and all around. *Trends in Endocrinology and Metabolism* , 25, 586-592. doi: 10.1016/j.tem.2014.06.011

Eyre-Walker, A., Keightley, P. D. (2007). The distribution of fitness effects of new mutations. *Nature Reviews Genetics* , 8:610-618. doi: https://doi.org/10.1038/nrg2146

Fan, H., & Chu, J-Y. (2007). A brief review of short tandem repeat mutation. *Genomics Proteomics Bioinformatics* , 5, 7-14. doi: 10.1016/S1672-0229(07)60009-6

Feder, J. L., Egan, S. P., & Nosil, P. (2012). The genomics of speciation-with-gene-flow. *Trends in Genetics* , 28, 342-350. doi: 10.1016/j.tig.2012.03.009

Gemayel, R., Vinces, M. D., Legendre, M., & Verstrepen, K. J. (2010). Variable tandem repeats accelerate evolution of coding and regulatory sequences. *Annual Review of Genetics* , 44, 445-477. doi: 10.1146/annurev-genet-072610-155046

Gemayel, R., Cho, J., Boeynaems, S., & Verstrepen, K. J. (2012). Beyond junk-variable tandem repeats as facilitators of rapid evolution of regulatory and coding sequences. *Genes* , 3, 461-480. doi: 10.3390/genes3030461

Gerhart-Hines, Z., Feng, D., Emmett, M. J., Everett, L. J., Loro, E., Briggs, E. R., . . . Lazar, M. A. (2013). The nuclear receptor Rev-erbα controls circadian thermogenic plasticity. *Nature* , 503, 410-413. doi: 10.1038/nature12642

Goldman, B. D. (2001). Mammalian photoperiodic system: formal properties and neuroendocrine mechanisms of photoperiodic time measurement.*Journal of Biological Rhythms* , 16, 283-301. doi: 10.1177/074873001129001980

Haerty, W., & Golding, G. M. (2010). Low-complexity sequences and single amino acid repeats: not just "junk" peptide sequences.*Genome* , 53, 753-762. doi: 10.1139/g10-063

Harrisson, K. A., Pavlova, A., Telonis-Scott, M., & Sunnucks, P. (2014). Using genomics to characterize evolutionary potential for conservation of wild populations. *Evolutionary Applications* , 7, 1008-1025. doi: https://doi.org/10.1111/eva.12149

Hedrick, P. W. (1999). Highly variable loci and their interpretation in evolution and conservation. *Evolution* , 53, 313-318. doi: 10.1111/j.1558-5646.1999.tb03767.x

Hemmer-Hansen, J., Nielsen, E. E., Frydenberg, J., & Loeschcke, V. (2007). Adaptive divergence in a high gene flow environment: Hsc70 variation in the European flounder (*Platichthys flesus L* .).*Heredity* , 99, 592-600. doi: 10.1038/sj.hdy.6801055

Hoving, C. L., Harrison, D. J., Krohn, W. B., Joseph, R. A., & O'Brien, M. (2005). Broad-scale predictors of Canada lynx occurrence in eastern North America. *Journal of Wildlife Management* , 69, 739-751. doi: https://doi.org/10.2193/0022-541X(2005)069[0739:BPOCLO]2.0.CO;2

Johnsen, A., Fidler, A. E., Kuhn, S., Carter, K. L., Hoffmann, A., Barr, I. R. . . . Kempenaers, B. (2007). Avian clock gene polymorphism: evidence for a latitudinal cline in allele frequencies. *Molecular Ecology* , 16, 4867-4880. doi: 10.1111/j.1365-294X.2007.03552.x

Johnson, W. E., & O'Brien, S. J. (1997). Phylogenetic reconstruction of the Felidae using 16S rRNA and NADH-5 mitochondrial genes. *Journal of Molecular Evolution* , 44, S98-S116. doi: 10.1007/pl00000060

Keenan, K., McGinnity, P., Cross, T. F., Crozier, W. W., & Prodöhl, P. A. (2013). diveRsity: an R package for the estimation and exploration of population genetics parameters and their associated errors.*Methods in Ecology and Evolution* , 4, 782-788. doi: https://doi.org/10.1111/2041-210X.12067

Keller, I., Taverna, A., & Seehausen, O. (2011). Evidence of neutral and adaptive genetic divergence between European trout populations sampled along altitudinal gradients. *Molecular Ecology* , 20, 1888-1904. doi: 10.1111/j.1365-294X.2011.05067.x

Kimura, M., & Crow, J. (1964). The number of alleles that can be maintained in a finite population. *Genetics* , 49, 725-738.

Koehler, G. M. (1990). Population and habitat characteristics of lynx and snowshoe hares in north central Washington. *Canadian Journal of Zoology* , 68, 845-851. doi: https://doi.org/10.1139/z90-122

Koen, E. L., Bowman, J., Lalor, J. L., & Wilson, P. J. (2014). Continental-scale assessment of the hybrid zone between bobcat and Canada lynx. *Biological Conservation* , 178, 107-115. doi: 10.1016/j.biocon.2014.07.016

Koen, E. L., Bowman, J., & Wilson, P. J. (2015). Isolation of peripheral populations of Canada lynx. *Canadian Journal of Zoology* , 530, 521-530. doi: https://doi.org/10.1139/cjz-2014-0227

Kolbe, J. A., & Squires, J. R. (2007). Circadian activity patterns of Canada lynx in western Montana. *Journal of Wildlife Management* , 71, 1607-1611. doi: 10.2193/2005-727

Kondratova, A. A., Dubrovsky, Y. V., Antoch, M. P., & Kondratov, R. V. (2010). Circadian clock proteins control adaptation to novel environment and memory formation. *Aging* , 2, 285-297. doi: 10.18632/aging.100142

Kozakiewicz, C. Z., Burridge, C. P., Funk, W. C., Salerno, P. E., Trumbo, D. R., Gagne, R. B., . . . (2019). Urbanization reduces genetic connectivity in bobcats (*Lynx rufus)* at both intra- and inter-population spatial scales. *Molecular Ecology* , 28, 5068-5085. doi:

Laidlaw, J., Gelfand, Y., Ng, K-W., Garner, H. R., Ranganathan, R., Benson, G., & Fondon, J. W. (2007). Elevated basal slippage mutation rates among the canidae. *Journal of Heredity* , 98, 452-460. doi: 10.1093/jhered/esm017

Laliberte, A. S., & Ripple, W. J. (2004). Range contractions of North American carnivores and ungulates. *BioScience* , 54, 123-138. doi: 10.1641/0006-3568(2004)054[0123:RCONAC]2.0.CO;2

Leadley, P., Pereira, H. M., Alkemade, R., Fernandez-Manjarrés, J. F., Proença, V., Scharlemann, J. P. W., & Walpole, M. J. (2010).*Biodiversity Scenarios: Projections of 21stcentury change in biodiversity and associated ecosystem services*(Secretariat of the Convention on Biological Diversity (SCBD) Technical Series no. 50). Montreal, ON: Canada.

Lenormand, T. (2002). Gene flow and the limits to natural selection.*Trends in Ecology and Evolution* , 17, 183-189. doi: https://doi.org/10.1016/S0169-5347(02)02497-7

Liedvogel, M., Szulkin, M., Knowles, S. C., Wood, M. J., & Sheldon, B. C. (2009). Phenotypic correlates of Clock gene variation in a wild blue tit population: evidence for a role in seasonal timing of reproduction.*Molecular Ecology* , 18, 2444-2456. doi: 10.1111/j.1365-294X.2009.04204.x

McKenny, D. W., Hutchinson, M. F., Papadopol, P., Lawrence, K., Pedlar, J., Campbell, K., . . . Owen, T. (2011). Customized spatial climate models for North America. *Bulletin of the American Meterological*

Society , 92, 1611-1622. doi:*https://doi.org/10.1175/2011BAMS3132.1*

Meirmans, P. G., & Hedrick, P. W. (2011). Assessing population structure: Fst and related measures. *Molecular Ecology Resources* , 11, 5-18. doi: https://doi.org/10.1111/j.1755-0998.2010.02927.x

Merilä, J., & Hendry, A. P. (2014). Climate change, adaptation, and phenotypic plasticity: the problem and the evidence. *Evolutionary Applications* , 7, 1-14. doi: https://doi.org/10.1111/eva.12137

Meröndun, J., Murray, D. L., & Shafer, A. B. A. (2019). Genome-scale sampling suggests cryptic epigenetic structuring and insular divergence in Canada lynx. *Molecular Ecology* , 28, 3186-3196. doi: https://doi.org/10.1111/mec.15131

Mesinger, F., DiMego, G., Kalnay, E., Mitchell, K., Shafran, P. C., Ebisuzaki, W., . . . Shi, W. (2006). North American Regional Reanalysis. *Bulletin of the American Meteorological Society* , 87, 343-360. doi:*https://doi.org/10.1175/BAMS-87-3-343*

Milligan, S. R., Holt, W. V., & Lloyd, R. (2009). Impacts of climate change and environmental factors on reproduction and development in wildlife. *Philosophical Transactions of the Royal Society B: Biology Sciences* , 364, 3313-3319. doi: https://doi.org/10.1098/rstb.2009.0175

Muhlfeld, C. C., Kovach, R. P., Jones, L. A., Al-Chokhachy, R., Boyer, M. C., Leary, R. F., . . . Allendorf, F. W. (2014). Invasive hybridization in a threatened species is accelerated by climate change.*Nature Climate Change Letters* , 4, 620-624. doi: 10.1038/nclimate2252

Nielsen, E. E., Hemmer-Hansen, J., Poulsen, N. A., Loeschcke, V., Moen, T., Johansen, T., . . . Carvalho, G. R. (2009). Genomic signatures of local directional selection in a high gene flow marine organism; the Atlantic cod (*Gadus morhua*). *BMC Evolutionary Biology* , 9, 276.

Økland, R. H. (1999). On the variation explained by ordination and constrained ordination axes. *Journal of Vegetation Science* , 10, 131-136. doi: https://doi.org/10.2307/3237168

Oksanen, J. (2012, January 25). Constrained ordination: tutorial with R and vegan [tutorial PDF]. Retrieved from https://www.yumpu.com/en/document/read/13318082/constrained-ordination-tutorial-with-r-and-vegan

Oksanen, J., Guillaume, F., Blanchet, M., Friendly, R., Kindt, P., Legendre, D., . . . Wagner, H. (2017). Vegan: Community ecology package. R package version 2.4-2.

O'Malley, K. G., Camara, M. D., & Banks, M. A. (2007). Candidate loci reveal genetic differentiation between temporally divergent migratory runs of Chinook salmon (*Oncorhynchus tshawytscha*).*Molecular Ecology* , 16, 4930-4941. doi: 10.1111/j.1365-294X.2007.03565.x

O'Malley, K. G., Ford, M. J., & Hard, J. J. (2010). Clock polymorphism in Pacific salmon: evidence for variable selection along a latitudinal gradient. *Proceedings of the Royal Society B: Biological Sciences* , 277, 3703-3714. doi: 10.1098/rspb.2010.0762

Parmesan, C. (2006). Ecological and evolutionary responses to recent climate change. *Annual Review of Ecology, Evolution, and Systematics* , 37, 637-669. doi:

https:doi.org/10.1146/annurev.ecolsys.37.091305.110100

Peers, M. J. L., Thornton, D. H., & Murray, D. L. (2012). Reconsidering the specialist-generalist paradigm in niche breadth dynamics: resource gradient selection by Canada lynx and bobcat. *PLoS One* , 7, e51488. doi: 10.1371/journal.pone.0051488

Peers, M. J. L., Thornton, D. H., & Murray, D. L. (2013). Evidence for large-scale effects of competition: niche displacement in Canada lynx and bobcat. *Proceedings of the Royal Society B: Biological Sciences* , 280, 20132495. doi: https://doi.org/10.1098/rspb.2013.2495

Poole, K. (2003). A review of the Canada lynx in Canada. *Canadian Field Naturalist*, 117, 360-376. doi: http://dx.doi.org/10.22621/cfn.v117i3.738

Prentice, M. B., Bowman, J., Lalor, J. L., McKay, M. M., Thomson, L. A., Watt, C. M., ... Wilson, P. J. (2017a). Signatures of selection in mammalian clock genes with coding trinucleotide repeats: implications for studying the genomics of high-pace adaptation. *Ecology and Evolution*, 7, 7254-7276. doi: https://doi.org/10.1002/ece3.3223

Prentice, M. B., Bowman, J., Khidas, K., Koen, E. L., Row, J. R., Murray, D. L., & Wilson, P. J. (2017b). Selection and drift influence genetic differentiation of insular Canada lynx (*Lynx canadensis*) on Newfoundland and Cape Breton Island. *Ecology and Evolution*, 7, 3281-3294. doi: https://doi.org/10.1002/ece3.2945

[dataset] Prentice, M. B., Bowman, J., Lalor, J. L., McKay, M. M., Thomson, L. A., Watt, C. M., ... Wilson, P. J.; 2017; Data from: Selection and drift influence genetic differentiation of insular Canada lynx (Lynx canadensis) on Newfoundland and Cape Breton Island; Dryad; doi: https://doi.org/10.5061/dryad.8ff46

Primmer, C. R., Papakostas, S., Leder, E. H., Davis, M. J., & Ragan, M. A. (2013). Annotated genes and nonannotated genomics: cross-species use of Gene Ontology in ecology and evolution research. *Molecular Ecology*, 22, 3216-3241. doi: 10.1111/mec.12309

Quinn, N. W. S., & Parker, G. (1987). Lynx. In: Novak, M., Baker, J. A., Obbard, M. E., & Malloch, B. (Eds.), *Wild furbearer management and conservation in North America* (pp. 683-694). North Bay, Ontario: Ontario Trappers Association.

R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available from .

Rellstab, C., Gugerli, F., Eckert, A. J., Hancock, A. M., & Holderegger, R. (2015). A practical guide to environmental association analysis in landscape genomics. *Molecular Ecology*, 24, 4348-4370. doi: https://doi.org/10.1111/mec.13322

Row, J. R., Gomez, C., Koen, E. L., Bowman, J., Murray, D. L., & Wilson, P. J. (2012). Dispersal promotes high gene flow among Canada lynx populations across mainland North America. *Conservation Genetics*, 13, 1259-1268. doi:10.1007/s10592-012-0369-3

Row, J. R., Wilson, P. J., Gomez, C., Koen, E. L., Bowman, J., Thornton, D., & Murray, D. L. (2014). The subtle role of climate change on population genetic structure in Canada lynx. *Global Change Biology*, 20, 2076-2086. doi: https://doi.org/10.1111/gcb.12526

Visser, M. E. (2008). Keeping up with a warming world: assessing the rate of adaptation to climate change. *Proceedings of the Royal Society B: Biological Sciences*, 275, 649-659. doi: https://doi.org/10.1098/rspb.2007.0997

Walpole, A. A., Bowman, J., Murray, D. L., & Wilson, P. J. (2012). Functional connectivity of lynx at their southern range periphery in Ontario, Canada. *Landscape Ecology*, 27, 761-773. doi: 10.1007/s10980-012-9728-1

Whitlock, M. C. (2000). Fixation of new alleles and the extinction of small populations: drift load, beneficial alleles, and sexual selection. *Evolution*, 54, 1855-1861. doi: 10.1554/0014-3820(2000)054[1855:FONAAT]2.0.CO;2

Wildlife Conservation Society (WCS) & Center for International Earth Science Information Network (CIESIN) Columbia University. (2005). *Last of the Wild Project* (version 2) (LWP-1): Global Human Influence Index (HII) Dataset (Geographic). Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). doi: http://dx.doi.org/10.7927/H4BP00QC. Accessed 27-10-2016

Willi, Y., Griffin, P., & Van Buskirk, J. (2012). Drift load in populations of small size and low density. *Heredity*, 110, 296-302. doi: 10.1038/hdy.2012.86

DATA ACCESSIBILITY

Neutral microsatellite data are available from the Dryad Digital Repository: .

Data for the *NR1D1* locus will be made available on a public repository pending revision/ acceptance of this article.

AUTHORS CONTRIBUTIONS

MBP, JB & PJW conceived the ideas and designed the methodology.

JB, KK, DLM & PJW facilitated sample collection.

MBP performed the laboratory work, analyzed the data and led the writing of the manuscript.

All authors contributed critically to the drafts and gave final approval for publication.

Table 1. Sample size and location of sampling of Canada lynx (*Lynx canadensis*), bobcat (*Lynx rufus*) and introgressed lynx-bobcat individuals across North America. Sample sizes are provided for datasets of 14 neutral microsatellite markers, the *NR1D1* locus, and the combined data. Samples missing >2 neutral loci were omitted.

Species	Province/State	Sample Size (Neutral markers)	Sample Size (*NR1D1*)	Total Sample Size (Combined data)
Canada lynx	Alaska	86	87	89
	Yukon	24	27	28
	British Columbia	190	193	196
	Alberta	104	109	110
	Manitoba	132	153	154
	Ontario	654	738	747
	Quebec	460	459	467
	Labrador	19	21	21
	New Brunswick	14	15	15
	Cape Breton Island	39	42	46
	Newfoundland	26	25	26
	Total	1,748	1,869	1,899
Bobcat	Oregon	0	1	1
	Washington	10	12	12
	British Columbia	63	66	67
	Idaho	59	63	67
	Montana	66	68	69
	Utah	1	1	1
	Wyoming	5	5	5
	Manitoba	11	12	12
	Texas	1	1	1
	Oklahoma	4	5	5
	Minnesota	78	96	97
	Ontario	20	20	20
	Wisconsin	5	5	5
	Michigan	50	51	53
	Tennessee	2	2	2
	Nova Scotia	90	92	93
	West Virginia	6	6	6

16

Species	Province/State	Sample Size (Neutral markers)	Sample Size (*NR1D1*)	Total Sample Size (Combined data)
	South Carolina	1	1	1
	Virginia	6	6	6
	New York	36	35	36
	Vermont	7	9	9
	Quebec	10	20	20
	Maine	42	43	45
	New Brunswick	102	106	117
	Total	675	726	750
Introgressed lynx	Saskatchewan	1	1	1
	Manitoba	1	1	1
	Total	2	2	2
Introgressed bobcats	Manitoba	1	0	1
	Texas	1	1	1
	New York	3	3	3
	Total	5	4	5

Table 2. Description of environmental variables used in RDA analyses of 1,899 Canada lynx (*Lynx canadensis*) sampled throughout Canada.

Variable	Description
Human influence	Influence index created by combining scores across 9 global data layers covering
Snow cover	National Centers for Environmental Prediction North American Regional Rean
Snow depth	
Mean temperature of the warmest quarter	Part of 19 bioclimatic variables estimated monthly and averaged annually for N
Mean temperature of the coldest quarter	
Precipitation of the wettest quarter	
Precipitation of the driest quarter	
Land cover	Global composites and land cover maps allocate land cover into 22 categories,

Table 3 . Results of a principal components analysis (PCA) on 9 environmental variables characterizing the environment of Canada lynx across their range in North America.

Principal Component	Proportion of Variance Explained (%)	Cumulative Proportion of Variance Explain
1	41.57	41.57
2	20.79	62.36
3	14.92	77.28
4	8.67	85.95

Table 4 . Eigenvectors of a principal components analysis (PCA) on 9 environmental variables characterizing the environment of Canada lynx across North America. The most influential variables on each component are represented in bold text.

Environmental variable	PC1	PC2	PC3	PC4
Mean snow cover	**-0.3655**	0.3392	-0.3302	-0.1584
Mean snow depth	-0.2931	0.3209	**-0.4326**	**-0.3993**
Mean temperature of the warmest quarter	0.1154	-0.3632	**-0.5576**	**0.4793**
Mean temperature of the coldest quarter	0.3431	-0.1637	0.3117	-0.3505
Precipitation of the wettest quarter	**0.3939**	-0.0795	-0.3289	**-0.3953**
Precipitation of the driest quarter	**0.4407**	-0.0423	-0.2107	**-0.3870**
Proportion of unsuitable lynx habitat	0.3072	**0.5531**	0.0832	0.1990
Proportion of suitable lynx habitat	-0.3031	**-0.5491**	-0.0907	-0.2131
Average human influence	0.3374	0.0883	-0.3579	0.2682

Table 5 . Results of RDA analyses testing for associations between neutral genetic structure, environmental variables, space, and functional genetic variation at the *NR1D1* locus in 3 subsets of data for Canada lynx: the full dataset, the "eastern" data subset, and the "western" data subset. The significance of the full and conditioned models are given by P-values (significant values shown in bold), and the total explainable inertia (variance) is given for each model. The proportion of inertia (variance) represents the proportion of total explainable variance in the full model that can be attributed to the effects of (1) purely spatial variables, (2) purely environmental variables, (3) purely neutral genetic structure and (4) the joint (collinear) effects of the spatial, environmental and neutral genetic structure datasets.

Model	Full dataset	Full dataset	Full dataset	"Eastern" subset of data	"Eastern" subset of data	"Eastern" subset of data	"Western" subset of data	"Western" subset of data	"Western" subset of data
	P-value	Inertia (variance)	Proportion of inertia (variance)	P-value	Inertia (variance)	Proportion of inertia (variance)	P-value	Inertia (variance)	Proportion of inertia (variance)
Full model NR1D1 ~ Environment + Space + Neutrals	**0.001**	0.1592	-	**0.001**	0.2128	-	0.140	0.2548	-

18

| Model	Full dataset	Full dataset	Full dataset	"Eastern" subset of data	"Eastern" subset of data	"Eastern" subset of data	"Western" subset of data	"Western" subset of data	"Western" subset of data
Conditioned model (pure spatial effects controlling for environment & neutral genetic structure) NR1D1 ~ Space + Condition(Environment + Neutrals)	0.808	0.0003	0.0019	**0.001**	0.0105	0.0493	0.533	0.0024	0.0
Conditioned model (pure environmental effects controlling for spatial & neutral genetic structure) NR1D1 ~ Environment + Condition(Space + Neutrals)	0.062	0.0033	0.0207	**0.008**	0.0065	0.0305	0.551	0.0052	0.0

Model	Full dataset	Full dataset	Full dataset	"Eastern" subset of data	"Eastern" subset of data	"Eastern" subset of data	"Western" subset of data	"Western" subset of data	"Western" subset of data
Conditioned model (pure neutral genetic structure effects controlling for environment & space) NR1D1 ~ Neutrals + Condition(Environment + Space)	**0.001**	0.1142	0.7173	**0.001**	0.1345	0.6320	0.154	0.2426	0.9
Joint environmental, spatial & neutral genetic structure effects		0.0414	0.2601		0.0613	0.2881		0.0046	0.0
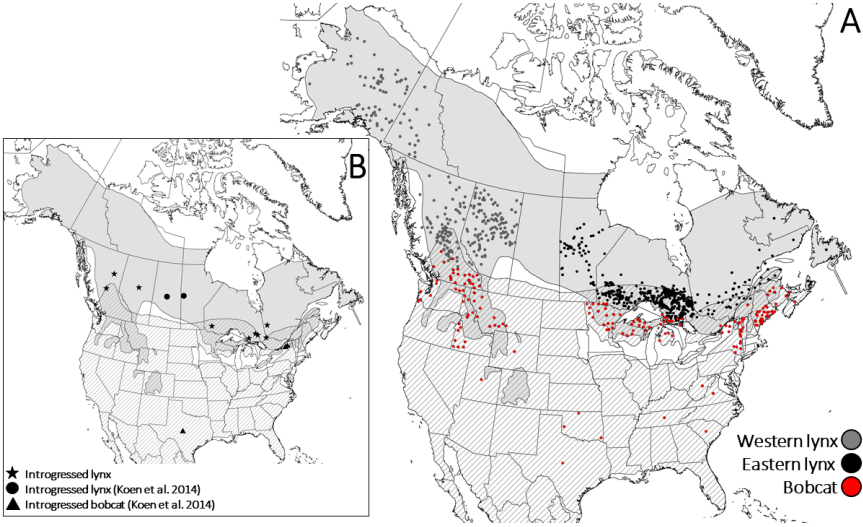
Figure 1 . (A) Samples of Canada lynx (*Lynx canadensis* , black and grey circles), bobcat (*Lynx rufus* , red circles), amplified at 14 neutral microsatellites and the *NR1D1* cTNR locus. Coordinates for lynx samples represent the centroid of sampling units (defined in methods). Samples of Canada lynx subdivided into "eastern" (black circles) and "western" (grey circles) lynx groups for RDA analyses. Coordinates of bobcat samples represent centroids of counties (USA), traplines (Canada), or province/state when finer resolution spatial information was not available. (B) Map of introgressed Canada lynx (*Lynx canadensis*) and bobcat (*Lynx rufus*) identified in Koen et al. (2014), and putative introgressed lynx identified via mixed genotypes at the *NR1D1* locus (current study). Samples from Saskatchewan, Alberta and Texas are plotted as the centroid of province/state. All other samples are plotted as centroids of traplines (Canada) or counties (USA) where the individual was trapped. Ranges of lynx and bobcat are represented by the solid grey and hatched grey areas, respectively. Online version in colour.
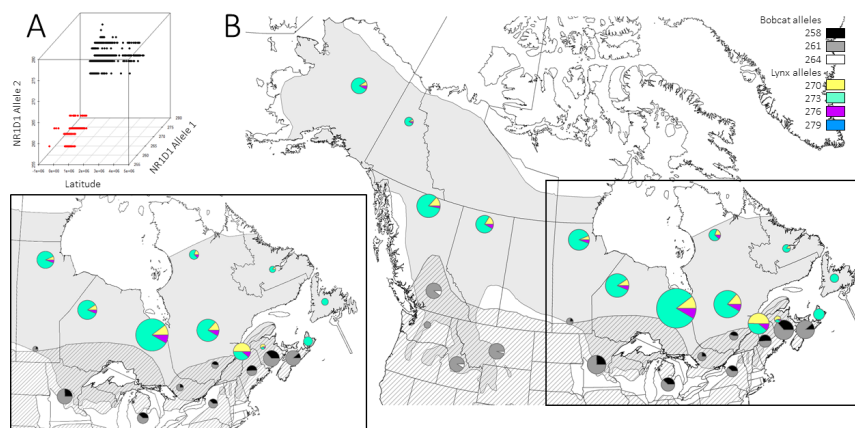
Figure 2. (A) Canada lynx (*Lynx Canadensis* , black points) and bobcat (*Lynx rufus,* red points) genotypes at the *NR1D1* locus plotted by latitude. (B) Allele frequency distribution of the *NR1D1* cTNR locus in Canada lynx and bobcat. The inset map shows a greater resolution of eastern Canada (outlined by the square on the full map). Canada lynx samples are characterized by the largest 4 alleles (shown in colour), and bobcat by the smallest 3 alleles (shown in black, white and grey). Size of the pie charts correspond to sample sizes. Bobcat sites with <10 samples were removed from this figure. Ranges of lynx and bobcat are represented by the solid grey and hatched grey areas, respectively. Online version in colour.
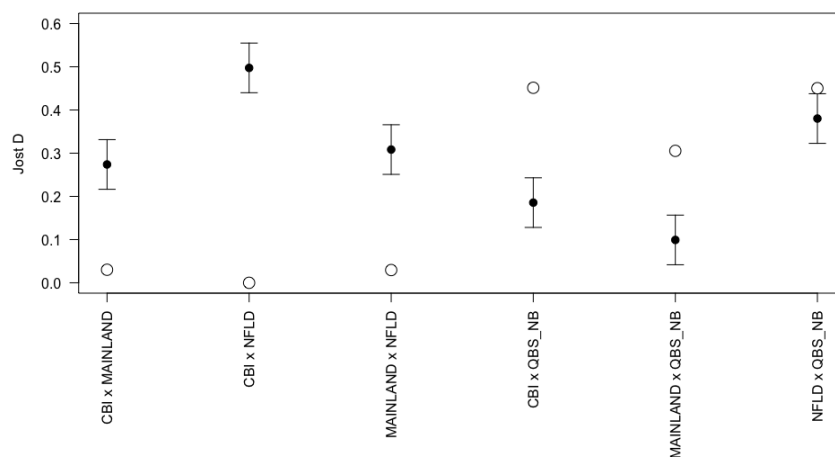
Figure 3. Pairwise point estimates of Jost's D for 14 neutral microsatellites (black points), with bars representing standard error of the mean, and point estimates of Jost's D for the *NR1D1* locus (white points) of 1,738 and 1,859 Canada lynx (*Lynx canadensis*) at neutral markers and the *NR1D1* locus, respectively. Site names are abbreviated and represent lynx from the four genetic clusters identified by STRUCTURE analyses: Cape Breton Island (CBI), Newfoundland (NFLD), lynx sampled from south of the St. Lawrence River in New Brunswick and Quebec (QBS_NB), and all remaining lynx sampled from continental Canada (MAINLAND).