Length polymorphisms at two candidate genes explain variation of migratory behaviors in blackpoll warblers (Setophaga striata)

Joel Ralston1 | Lydia Lorenc1 | Melissa Montes1 | William V. DeLuca2 | Jeremy J. Kirchman3 | Bradley K. Woodworth4,5 | Stuart A. Mackenzie6 | Amy Newman4 | Hilary A. Cooke7 | Nikole E. Freeman4 | Alex O. Sutton4 | Lila Tauzer7 | D. Ryan Norris4

Abstract
Migratory behaviors such as the timing and duration of migration are genetically inherited and can be under strong natural selection, yet we still know very little about the specific genes or molecular pathways that control these behaviors. Studies in candidate genes *Clock* and *Adcyap1* have revealed that both of these loci can be significantly correlated with migratory behaviors in birds, though observed relationships appear to vary across species. We investigated geographic genetic structure of *Clock* and *Adcyap1* in four populations of blackpoll warblers (*Setophaga striata*), a Neotropical–Nearctic migrant that exhibits geographic variation in migratory timing and duration across its boreal breeding distribution. Further, we used data on migratory timing and duration, obtained from light-level geolocator trackers to investigate candidate genotype–phenotype relationships at the individual level. While we found no geographic structure in either candidate gene, we did find evidence that candidate gene lengths are correlated with five of the six migratory traits. Maximum *Clock* allele length was significantly and negatively associated with spring arrival date. Minimum *Adcyap1* allele length was significantly and negatively associated with spring departure date and positively associated with fall arrival date at the wintering grounds. Additionally, we found a significant interaction between *Clock* and *Adcyap1* allele lengths on both spring and fall migratory duration. *Adcyap1* heterozygotes also had significantly shorter migration duration in both spring and fall compared to homozygotes. Our results support the growing body of evidence that *Clock* and *Adcyap1* allele lengths are correlated with migratory behaviors in birds.

**KEYWORDS**
*Adcyap1*, avian migration, *Clock*, phenology, pituitary adenylate cyclase-activating polypeptide
Avian migration involves a suite of complex behavioral and physiological responses to exogenous seasonal cues including changes in diet and metabolism, the onset of migratory restlessness (zugunruhe), and precise directional orientation (Bairlein, Eikenaar, & Schmaljohann, 2015; Berthold, 1996; Gwinner, 1990; Jenni & Schaub, 2003). The culmination of these responses allow individual birds to accomplish amazing migratory feats, sometimes flying over ocean for several days at a time or navigating thousands of kilometers through unfamiliar habitat between wintering and breeding grounds (Bairlein et al., 2012; Battley et al., 2012; DeLuca et al., 2019, 2015; Gill et al., 2009; McKinnon, Artuso, & Love, 2017). Each of these behavioral and physiological responses is likely genetically inherited (Berthold & Pulido, 1994; Helbig, 1991; Liedvogel & Lundberg, 2014; Pulido, Berthold, Mohr, & Querner, 2001) and could be under natural selection (Gienapp, Leimu, & Merila, 2019, 2015; Gill et al., 2009; McKinnon, Artuso, & Love, 2017). Yet we still know very little about the specific genes or molecular pathways that control these behaviors (Liedvogel, Åkesson, & Bensch, 2011). Further, it is unknown whether molecular pathways controlling migration vary among animal taxa, or even among migratory bird taxa given that migration is evolutionarily labile and likely arose independently multiple times within birds (Pulido, 2007; Zink, 2011). Investigating the genetic control of migratory behaviors in a diversity of bird taxa will therefore allow us to better understand the evolution of migration within birds, as well as ongoing natural selection on these behaviors in response to environmental change (Pulido & Berthold, 2010).

Recent studies have attempted to identify candidate genes that may be linked to migration in insects, birds, and other vertebrates (Contina, Bridge, & Kelly, 2016; Delmore et al., 2015; Johnston, Paxton, Moore, Wayne, & Smith, 2016; Lemopoulos, Uusi-Helkkilä, Huusko, Vasmägi, & Vainikka, 2018; Lundberg et al., 2013, 2017; Merlin & Liedvogel, 2019; Mueller, Pulido, & Kempenaers, 2011; Steinmeyer, Mueller, & Kempenaers, 2009; Zhu, Gegear, Casselman, Kanginakudru, & Reppert, 2009). Two genes that have received considerable attention, especially in regard to migratory phenology, are Clock and Adcyap1. Clock plays a central role in regulating the circadian oscillator gene complex, which controls circadian and circannual rhythmicity in response to exogenous cues (Panda, Hogensche, & Kay, 2002; Yu & Hardin, 2006). Adcyap1 codes for pituitary adenylate cyclase-activating polypeptide (PACAP) which influences circadian rhythms in part by directly activating Clock and other genes in the circadian oscillator complex (Nagy & Csernus, 2007). PACAP also has multiple effects on physiology and behavior throughout the body (Vaudry et al., 2009), including melatonin production in the pineal gland (Csernus et al., 2004) and a role in processing light signals from the retina into neuronal signals (Hannibal et al., 1997), both of which likely play a role in the photoperiodic control of seasonality in birds (Dawson, King, Bentley, & Ball, 2001). Clock and Adcyap1, therefore, represent two candidate genes that natural selection may act on independently or in concert to shape migratory behaviors in natural populations of migratory organisms.

Several studies have shown a correlation between length polymorphisms in Clock and Adcyap1 and migratory traits in birds (Table 1). For Clock, individuals with a greater number of glutamine repeats in the 3’ polyglutamine tail (i.e., longer alleles) exhibit delayed migratory and breeding phenology (Bazzi et al., 2015; Bourret & Garant, 2015; Caprioli et al., 2012; Liedvogel, Szulkin, & Knowles, 2009; Saino et al., 2015) and longer migratory distance (Peterson et al., 2013), relative to individuals with shorter alleles. For Adcyap1, longer alleles have been shown to be associated with greater migratory restlessness (Mueller et al., 2011; Peterson et al., 2013), earlier spring arrival dates (Mettler, Segelbacher, & Schaefer, 2015), and earlier postnatal dispersal (Chakarov, Jonker, Boerner, Hoffman, & Krüger, 2013). Additionally, in one species, more northerly breeding populations had longer Adcyap1 alleles on average than southerly populations (Bazzi et al., 2016), which may reflect geographic variation in migratory strategies or phenological schedules resulting from local adaptation to environmental cues (Johnsen et al., 2007). However, the relationships between candidate genes and migratory phenotypes may also be influenced by local environmental factors such as temperature, photoperiod, and breeding density (Bourret & Garant, 2015), potentially complicating interpretations of geographic patterns within species. Further complicating the study of candidate genes and migratory traits, an increasing number of studies in birds have found no correlation between Clock or Adcyap1 allele length and migratory behavior (Table 1; Bazzi et al., 2016, 2017; Contina, Bridge, Ross, Shipley, & Kelly, 2018; Dor et al., 2012). The interspecific variation in genotype–phenotype correlations for these candidate migratory-phenology genes highlights the challenge thus far of generalizing the expected relationship between length polymorphism and migration timing, as well as identifying the mechanism by which length variants affect migratory behaviors.

Given the potential evolutionary and ecological importance of these migratory candidate genes, and the variation in observed patterns across species, it is valuable to continue to build evidence to either support or refute a role of these genes in migratory behavior in different avian species, particularly to help understand why patterns vary across species. Here, we contribute to this growing number of studies by examining relationship between Adcyap1 and Clock and migratory behavior in the blackpoll warbler (Setophaga striata), a small (12 g) long distance Neotropical–Nearctic migrant (Figure 1) with geographic variation in migratory behaviors (DeLuca, Holberton, Hunt, & Eliason, 2013; DeLuca et al., 2015; Morris, Covino, Jacobs, & Taylor, 2016).

Blackpoll warblers breed across the North American boreal forest from Alaska to Newfoundland and in isolated montane fir forests at the southern periphery of their range in northeastern United States (Figure 2; DeLuca et al., 2013). Annual blackpoll warbler migration between the northern breeding grounds and wintering grounds in the Amazon basin of northern South America is one of the longest migratory routes for any North American songbird (DeLuca et al., 2013; Morris et al., 2016), and was recently tracked from geographically distant breeding populations using light-level geolocators (DeLuca et al., 2015, 2019; Figure 1). Birds depart the wintering grounds in mid-April to mid-May and travel through the Greater Antilles and continental United States
to their northern breeding territories, a trip that can vary in duration considerably depending on breeding location (mean 34 days, range 17–49 days; DeLuca et al., 2019). After departing the breeding grounds in August through October, blackpoll warblers first migrate southeastward to the Atlantic coast of United States and Canada, then migrate south over the Atlantic Ocean to their South American wintering grounds, a nonstop trans-oceanic flight of up to 3,000 km that may take up to three days (DeLuca et al., 2015, 2019).

The large breeding range means that blackpoll warblers from different breeding localities vary in the timing, direction, distance, and duration of their migrations. Birds that breed further north and west migrate further and take longer; these populations tend to depart the wintering grounds earlier but arrive on the breeding grounds at

### TABLE 1
**Studies of association between Clock and Adcyap1 allele lengths, migratory behaviors, and breeding latitude in birds**

<table>
<thead>
<tr>
<th>Gene</th>
<th>Study</th>
<th>Species</th>
<th>Phenology</th>
<th>Migratory propensity</th>
<th>Breeding latitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clock</td>
<td>Johnsen et al. (2007)</td>
<td>Bluethroat, <em>Luscinia svecica</em></td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Blue tit, <em>Cyanistes caerules</em></td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Liedvogel et al. (2009)</td>
<td>Blue tit, <em>Cyanistes caerules</em></td>
<td>+/0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Dor et al. (2011)</td>
<td>Barn swallow, <em>Hirundo rustica</em></td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mueller et al. (2011)</td>
<td>Eurasian blackcaps, <em>Sylvia atricapilla</em></td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Caprioli et al. (2012)</td>
<td>Barn swallow, <em>Hirundo rustica</em></td>
<td>+/0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Dor et al. (2012)</td>
<td>Swallows, <em>Tachycineta</em></td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Chakarov et al. (2013)</td>
<td>Common buzzard, <em>Buteo buteo</em></td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Peterson et al. (2013)</td>
<td>Dark-eyed junco, <em>Junco hyemalis</em></td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Kuhn et al. (2013)</td>
<td>Pied flycatcher, <em>Ficedula hypoleuca</em></td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bazzi et al. (2015)</td>
<td>Barn swallow, <em>Hirundo rustica</em></td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bourret and Garant (2015)</td>
<td>Tree swallow, <em>Tachycineta bicolor</em></td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pied flycatcher, <em>Ficedula hypoleuca</em></td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tree pipit, <em>Anthus trivialis</em></td>
<td>+/0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Winchat, <em>Saxicola ruberta</em></td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bazzi et al. (2016)</td>
<td>WILSON’S warbler, <em>Cardellina pusilla</em></td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bazzi et al. (2017)</td>
<td>Willow warbler, <em>Phylloscopus trochilus</em></td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Contina et al. (2018)</td>
<td>Painted bunting, <em>Passerina ciris</em></td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Romano et al. (2018)</td>
<td>Yellow-legged gull, <em>Larus michahellis</em></td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adcyap1</td>
<td>Mueller et al. (2011)</td>
<td>Eurasian blackcaps, <em>Sylvia atricapilla</em></td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Chakarov et al. (2013)</td>
<td>Common buzzard, <em>Buteo buteo</em></td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Peterson et al. (2013)</td>
<td>Dark-eyed junco, <em>Junco hyemalis</em></td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bourret and Garant (2015)</td>
<td>Tree swallow, <em>Tachycineta bicolor</em></td>
<td>-/+</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mettlер et al. (2015)</td>
<td>Eurasian blackcaps, <em>Sylvia atricapilla</em></td>
<td>-/0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Saino et al. (2015)</td>
<td>Nightingale, <em>Luscinia megarhynchos</em></td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pied flycatcher, <em>Ficedula hypoleuca</em></td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tree pipit, <em>Anthus trivialis</em></td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Winchat, <em>Saxicola ruberta</em></td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bazzi et al. (2016)</td>
<td>Wilson’s warbler, <em>Cardellina pusilla</em></td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bazzi et al. (2017)</td>
<td>Willow warbler, <em>Phylloscopus trochilus</em></td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Contina et al. (2018)</td>
<td>Painted bunting, <em>Passerina ciris</em></td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Romano et al. (2018)</td>
<td>Yellow-legged gull, <em>Larus michahellis</em></td>
<td>0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Migratory propensity includes studies of migratory restlessness, distance, and duration. Symbols indicate positive (+), negative (−), or no observed relationships (0).

*Relationships that varied by sex (F/M).

*Relationships that varied by latitude (low latitude/high latitude).
similar times compared to birds breeding further east (DeLuca et al., 2015, 2019). Like in the spring, fall migration timing and duration varies across breeding populations with western breeding birds departing earlier and taking longer to arrive at wintering grounds compared to eastern breeding birds (DeLuca et al., 2015, 2019).

We assessed variation in Clock and Adcyap1 across four breeding populations of blackpoll warbler and tested for correlations between allele lengths and the timing of departure, arrival, and duration in both spring and fall migratory periods. Based on published evidence from other species (Table 1), we predicted allele lengths in both Clock and Adcyap1 would be positively associated with migratory distance and duration. We predicted Clock would have a positive relationship with the timing of migration (departure and arrival dates), and Adcyap1 would have a negative relationship with migratory timing.

2 | MATERIALS AND METHODS

2.1 | Genetic sampling

We collected Blackpoll Warbler blood and tissue samples from throughout the breeding range using mist nets during the May–July breeding seasons from 2009 to 2019 (n = 64). Blood samples taken in the field were collected using brachial venipuncture and were stored in lysis buffer at ambient temperature until they were transported to a genetics laboratory where they were stored at −20°C. Samples collected in the field were supplemented with a small number of frozen tissues from natural history museum collections, also collected from the breeding seasons from 2008 to 2009 (n = 8, Appendix 1). Our total sample (n = 72) included only breeding adults, no nestlings, or known relatives, and was mostly male (64/72, 89% males). We extracted DNA from all samples using a DNeasy Blood and Tissue Extraction Kit (Qiagen, Valencia, CA) following the manufacturer’s protocol with one modification: Final elution was conducted in two rounds of 75 μL. All blood samples and DNA extracts used in the current study, including those for geolocator tagged birds, are archived at Saint Mary’s College, Notre Dame, Indiana, USA, or at New York State Museum, Albany, New York, USA.

We amplified Clock and Adcyap1 loci in 10 μl polymerase chain reactions (PCRs) using a Qiagen Multiplex PCR Master Mix (HotStar Taq DNA Polymerase, dNTP mix, final concentration of 3.0 mM MgCl₂), 0.2 μM reverse and fluorescent-labeled forward primers (Applied Biosystems, Foster City, CA), and <200 ng of template DNA. Primer sequences were obtained from Johnsen et al. (2007) and Steinmeyer et al. (2009). Our PCR thermal regime included a 15-min initial denaturation at 95°C; 35 cycles of 30-s denaturation at 94°C, 90-s primer annealing at 54°C, and 90-s extension at 72°C; and a final extension period of 10 min at 72°C. Final PCR products were mixed with LIZ500 size standard (Thermo Fisher Scientific), diluted in Hi-Di formamide, and sent to the University of Notre Dame Genomics and Bioinformatics Core Facility for genotyping on an Applied Biosystems 3730xl Genetic Analyzer. We used Peak Scanner (Applied Biosystems) to examine electropherograms and estimate the sizes of the alleles (measurement unit = number of base pairs) from each individual at Clock and Adcyap1 loci.

2.2 | Analysis of geographic variation

We grouped samples into four populations for analysis: Alaska (n = 14), Yukon (n = 11), Manitoba (n = 18), and Eastern (n = 29). The Alaska population included samples from Nome (n = 10), Denali National Park (n = 2), and southwest Alaska (n = 2). Four sampling locations were grouped together into the Eastern population: Adirondack and Catskill Mountains, New York (n = 8); Cape Breton, Nova Scotia (n = 11); western Newfoundland (n = 5); and southeastern Labrador (n = 5). All samples from Yukon are from Whitehorse, Yukon Territories, and all samples from Manitoba are from Churchill, Manitoba. Previous analysis of neutral loci revealed significant genetic structure in mitochondrial DNA marker between Newfoundland and other Eastern subpopulations (Ralston & Kirchman, 2012), but no significant structure and a large number of shared alleles at neutral microsatellite markers (Ralston & Kirchman, 2013). Due to the small sample size of Newfoundland in the current study, we group these samples within the Eastern population.

We used the package hierfstat (Goudet, 2004) in program R version 3.5.5 (R Core Team, 2019) to calculate overall FST and FIS. We used the program ARLEQUIN version 3.5 (Excoffier & Lischer, 2010) to test for deviation from Hardy–Weinberg equilibrium, evidence of genetic linkage disequilibrium between the two loci, and to assess genetic population differentiation using pairwise FST and an analysis of molecular variance (AMOVA). Significance for each test was determined using default ARLEQUIN settings for number of permutations and number of steps in Markov chain Monte Carlo (MCMC) algorithms.

We further investigated geographic structure at these loci using the program STRUCTURE version 2.3.4 (Pritchard, Stephens, & Donnelly, 2000), which uses a Bayesian clustering analysis to determine the most likely number of populations (K). We used an admixture model with an assumption of correlated allele frequencies among populations, and population of each sample as a prior. We ran 10 iterations for each K ranging from 1 to 4 with 1,000,000 MCMC steps following a burn-in of 100,000 steps, and used mean natural log probability to determine the most likely number of populations. We used the program DISTRACT (Rosenberg, 2004) to visualize population assignment of each individual from the STRUCTURE run with the greatest log probability for each value of K. Additionally, we performed a principal components analysis (PCA) implemented in the R package ade4 (Dray & Dufour, 2007). We visualized clustering of genotypes in ordination space using principal components with eigenvalues greater than 1.0.

While the above genetic analyses assess variation in the frequencies of alleles across populations, they do not directly assess patterns in allele length per se. In other words, in the above analyses, alleles are treated as categorical, while we are also interested in...
allele length varied across either sampling latitude or longitude. It is unknown whether there are any allelic interactions or dominance effects at these loci relative to the migratory traits of interest, though previous studies have suggested dominance of the longer allele in Clock (Saino et al., 2015) and other genes with 3′ polyglutamine repeats (Ross, 2002). Due to small sample sizes per genotype, and high variability in Adcyap1, we could not directly compare genotypes for dominance effects (as in Liedvogel et al., 2009; Saino et al., 2015). However, we did run separate GLMs using either minimum allele, maximum allele, or mean allele length for each individual as the dependent variable. If longer Clock alleles are dominant, we would expect to see stronger relationships between individuals’ maximum allele length and migratory traits.

2.3 | Analysis of candidate genes relative to individual migratory traits

To investigate correlations between candidate genes and migratory behaviors at the individual level, we combined genotype data described above with migratory data obtained from light-level geolocators. In a separate study specifically addressing migratory routes and phenology, DeLuca et al. (2019) tracked migration in blackpoll warblers breeding in Alaska, Yukon, and Manitoba with Biotrack model MK-6 light-level geolocators (Wareham, UK), and took blood samples from tagged individuals for the current study (Alaska n = 5, Yukon n = 4, Manitoba n = 8). All individuals used in this analysis were unrelated breeding adult males. Full details on the methods of geolocator deployment and the analysis of geolocator track data are available in DeLuca et al. (2019). For the current study, we analyze six migratory traits related to phenotype and duration of migration: (a) spring departure date from wintering grounds, (b) duration of spring migration, (c) spring arrival date on the breeding grounds, (d) fall departure date from breeding grounds, (e) duration of fall migration, and (f) fall arrival date on the wintering grounds.

We tested for correlations between Clock and Adcyap1 allele lengths and each of the six migratory traits using GLMs. To account for known variation in migratory traits across populations (DeLuca et al., 2019), we include population, coded as a factor, as a predictor variable in the models. Like in the population-level analyses, we built separate GLMs using either minimum allele, maximum allele, or mean allele length for each individual. For six individuals, geolocator track information was not available from the spring migration.
period. Therefore, analyses comparing spring departure date, spring duration, or spring arrival date with allele length had lower sample sizes ($n = 12, n = 12, n = 11$, respectively; $n = 17$ for all fall migratory traits). We repeated these analyses using longitude and latitude as covariables instead of population to determine whether our results were robust to this selection.

Next, we tested for an interaction between Clock and Adcyap1 on each of the six migratory traits using GLMs with population, Clock allele length, Adcyap1 allele length, and a Clock × Adcyap1 interaction term as the independent variables. So as to not vastly increase our number of tests by comparing all possible allele length combinations of Clock and Adcyap1, we only tested GLMs with interactions between allele lengths that showed significant or marginally significant correlations in the individual gene models described above (see Section 3). If we found a significant Clock × Adcyap1 interaction effect on a migratory trait, we do not further discuss single gene effects for this migratory trait.

Lastly, we compared each of the six migratory traits between homozygous and heterozygous individuals for both Clock and Adcyap1 using GLMs with location and heterozygosity as the predictor variables. If heterozygosity correlates with individual fitness and migratory performance (Mettler et al., 2015), we would expect heterozygotes to have significantly shorter migratory durations and earlier arrival times. While several studies have shown sex-specific effects of candidate genes on phenology (e.g., Mettler et al., 2015; Bazzi et al., 2017), we could not test for sex effects because all 17 geolocator birds were male.

3 | RESULTS

We successfully genotyped 72 Blackpoll Warblers at candidate genes Clock and Adcyap1. We observed four Clock alleles ranging in size from 186 to 195 base pairs (Table 2). We found relatively greater variation in Adcyap1: we detected 13 Adcyap1 alleles ranging from 153 to 169 base pairs, of which five were found in single individuals (Table 2). Observed heterozygosity was fairly high ($H_e = 0.722$ for each locus), with only five individuals being homozygous at both loci. None of the populations deviated from Hardy–Weinberg equilibrium (all $p > 0.05$, Table 3), and we found no evidence of linkage disequilibrium ($p > 0.05$).

We found no evidence of structured geographic variation in either Clock or Adcyap1. Overall $F_{ST}$ was $-0.003$, and $F_{IS}$ was 0.070. All pairwise $F_{ST}$ values between populations were nonsignificant (Table 3), and the AMOVA revealed that only 0.72% of the variation in candidate genes was explained by difference among populations, while 99.28% corresponded to variation within populations. The STRUCTURE analysis, similarly, suggested no geographic structure with $K = 1$ as the most likely number of groups (Figure 3a). For all values of $K$, the assignment probability of each individual was roughly equal for all populations (i.e., the probability of any individual belonging to any of $K$ populations $= 1/K$; Figure 3b). The first two principal components from a PCA on genotype data had eigenvalues greater than 1.0 (1.61 and 1.26, respectively) and together explained 71.68% of the variation in genotype data. A PCA plot on these axes revealed no clustering by population. All populations overlapped in ordination space (Figure 3c). We found no significant relationship between allele length for either locus and either longitude or latitude, regardless of whether individual minimum, maximum, or mean allele lengths were considered (all $p > 0.05$, all $R^2$ values <0.05; Appendix 2). From our GLMs, we found a significant correlation between population and spring and fall departure, and spring and fall duration, but not with spring and fall arrival. These results are not explored further here, as population variation in migratory timing is more thoroughly discussed in DeLuca et al. (2019).

Despite finding no evidence of geographic variation in candidate genes, we did find evidence that both Clock and Adcyap1 allele length were correlated with migratory behaviors at an individual level (Table 4). Maximum Clock allele length was significantly and negatively associated with spring arrival date on the breeding ground (GLM: $F_{3,7} = 9.466, p = 0.007, R^2 = 0.718; \beta_{\text{MaxClock}} = -1.25, P_{\text{MaxClock}} = 0.024$); individuals with longer Clock alleles tended to arrive to the breeding territories earlier in the spring (Figure 4). We found no other correlations for Clock allele length and the other five migration traits from our single gene models (Appendix 3). We found significant correlation between individual’s minimum Adcyap1 allele length and the timing of spring and fall migration. Blackpoll warblers with longer minimum Adcyap1 alleles departed earlier in the spring ($F_{3,8} = 6.936, p = 0.013, R^2 = 0.618; \beta_{\text{MinAdcyap1}} = -5.615, P_{\text{MinAdcyap1}} = 0.007$) and arrived later to the southern wintering grounds ($F_{3,13} = 2.418, p = 0.113, R^2 = 0.210; \beta_{\text{MinAdcyap1}} = 1.212, P_{\text{MinAdcyap1}} = 0.043$; Figure 5). A single individual that was homozygous for the longest minimum Adcyap1 allele had by far the earliest spring departure, the longest spring and fall duration, and the latest fall arrival. Removal of this individual from GLM analyses resulted in nonsignificant relationships between minimum Adcyap1 allele length and both spring departure ($F_{3,9} = 8.531, p = 0.010, R^2 = 0.693; \beta_{\text{MinAdcyap1}} = -3.239, P_{\text{MinAdcyap1}} = 0.088$) and fall arrival ($F_{3,12} = 2.469, p = 0.112, R^2 = 0.227; \beta_{\text{MinAdcyap1}} = 0.580, P_{\text{MinAdcyap1}} = 0.535$). The correlation between Adcyap1 and fall arrival was also significant when considering individuals’ mean allele length (Table 4). Results were consistent when using longitude and latitude as covariates instead of population.

We also found an interaction effect between Clock and Adcyap1 allele lengths on migratory duration (Figure 6). When we used mean Clock allele length and minimum Adcyap1 allele length, this interaction was significantly correlated with both spring duration ($F_{5,6} = 10.15, p = 0.007, R^2 = 0.806; \beta_{\text{MeanClock} \times \text{MinAdcyap1}} = 0.019$) and fall duration ($F_{5,11} = 22.38, p < 0.001, R^2 = 0.870; \beta_{\text{MeanClock} \times \text{MinAdcyap1}} = 0.007$). For individuals with shorter-than-average mean Clock allele lengths, both spring duration and fall duration were determined by a significant positive relationship with minimum Adcyap1 allele length. For individuals with longer-than-average mean Clock allele lengths, duration was not correlated with Adcyap1 length (Figure 6). When the individual with the longest minimum Adcyap1
TABLE 2  Allele frequencies for two candidate genes, Clock and Adcyap1, at four populations of blackpoll warblers

<table>
<thead>
<tr>
<th>Population</th>
<th>Clock allele lengths</th>
<th>Adcyap1 allele lengths</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>186</td>
<td>189 192 195 153 155</td>
</tr>
<tr>
<td>Alaska</td>
<td>5</td>
<td>12 4 7 0 1 4 13 0 4 0</td>
</tr>
<tr>
<td>Yukon</td>
<td>0</td>
<td>11 5 6 1 2 6 7 0 4 0</td>
</tr>
<tr>
<td>Manitoba</td>
<td>4</td>
<td>13 13 6 0 3 6 14 0 4 0</td>
</tr>
<tr>
<td>Eastern</td>
<td>6</td>
<td>20 15 17 0 1 10 17 1 19 1</td>
</tr>
</tbody>
</table>

Note: Allele numbers represent lengths in number of base pairs.

allele is removed from the analyses, this relationship becomes non-
significant for spring duration ($F_{5,5} = 5.538, p = 0.042, R^2 = 0.694$; $p_{\text{MeanClock} \times \text{MinAdcyap1}} = 0.120$), but remains significant for fall duration ($F_{5,10} = 23.37, p < 0.001, R^2 = 0.882$; $p_{\text{MeanClock} \times \text{MinAdcyap1}} = 0.011$). This Clock × Adcyap1 interaction effect was weaker when using either maximum Clock allele length or mean Adcyap1 allele length (Table 5).

Lastly, we tested for differences in each of the migratory traits between homozygotes and heterozygotes of each locus. We found individuals that were heterozygous at Adcyap1 had significantly shorter spring duration after accounting for population origin in a GLM (mean homozygote residuals = 8.47; mean heterozygote residuals = −8.47, $p = 0.030$) and significantly shorter fall duration (mean homozygote residuals = 3.26; mean heterozygote residuals = −2.90, $p = 0.031$) compared to homozygotes.

4 | DISCUSSION

We found evidence suggesting allele lengths in candidate genes Clock and Adcyap1 are correlated with spring and fall migratory behavior in blackpoll warblers. Despite no geographic variation across populations in allele frequencies or lengths, variation in candidate gene allele lengths was associated with nearly every migratory trait analyzed at the individual level. It is worth noting that we conducted a large number of statistical tests ($N = 78$ GLMs), increasing the chance of type I errors. However, with $\alpha = 0.05$, we would expect an average of 3.7 of the 78 tests to yield significant results at random. We find significant correlations ($p < 0.05$) in 12 cases, significantly more than expected at random (binomial test, $p < 0.001$). Given that the observed relationships are mostly in agreement with predicted direction of relationships, and are consistent with studies in other species, it is likely that our results represent real relationships between the variables and not spurious statistical effects. While sample sizes are low given the low recovery rates of geolocator-fitted birds (DeLuca et al., 2019), future work that aims to increase sample sizes may allow more rigorous evaluation of links between candidate genes and migration behavior. For example, with more birds, analyses could control for other factors known to influence migratory timing and speed, such as age, sex, or environmental conditions experienced during and leading up to migration, which could obscure genetic contributions to migration phenology.

The only consistent effect we found for Clock alleles was a negative correlation with spring arrival date (Figure 4, Table 4). This was a surprising result as all other published studies that report a significant effect of Clock allele length on spring phenology in birds report a positive correlation (Table 1; Bazzi et al., 2015; Bourret & Garant, 2015; Caprioli et al., 2012; Liedvogel et al., 2009; Saino et al., 2015). Why blackpoll warblers would exhibit an opposite pattern at this locus is unclear, though this highlights the emerging pattern of variation across species in migratory genotype–phenotype relationships and a need for further detailed studies that investigate the functional influence of candidate gene alleles. Previous studies have suggested that longer Clock alleles are dominant over shorter alleles in their influence on phenology (Bazzi et al., 2015; Saino et al., 2015). Our results seem to support this given that we find a stronger relationship between maximum Clock allele length and spring arrival time, compared to minimum Clock alleles.

TABLE 3  Summary statistics for each candidate gene, Clock and Adcyap1, for four populations of blackpoll warbler

<table>
<thead>
<tr>
<th>Population</th>
<th>n</th>
<th>Clock</th>
<th>Adcyap1</th>
<th>Pairwise $F_{ST}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$H_o$</td>
<td>$H_e$</td>
<td>A</td>
</tr>
<tr>
<td>Alaska</td>
<td>14</td>
<td>0.571</td>
<td>0.728</td>
<td>4</td>
</tr>
<tr>
<td>Yukon</td>
<td>11</td>
<td>0.636</td>
<td>0.654</td>
<td>3</td>
</tr>
<tr>
<td>Manitoba</td>
<td>18</td>
<td>0.667</td>
<td>0.719</td>
<td>4</td>
</tr>
<tr>
<td>Eastern</td>
<td>29</td>
<td>0.862</td>
<td>0.730</td>
<td>4</td>
</tr>
</tbody>
</table>

Note: Sample size (n), observed ($H_o$) and expected heterozygosity ($H_e$), number of alleles (A) per locus, and pairwise $F_{ST}$. None of the $H_o$ were statistically different from $H_e$, and no pairwise $F_{ST}$ were significant.
Our results showed longer minimum Adcyap1 alleles were associated with earlier spring departure and later fall arrival on the wintering grounds. These results are generally consistent with work previously published on other species (Table 1). While no other studies to our knowledge have shown a significant correlation between Adcyap1 and arrival date on the wintering grounds as we show here, these results are consistent with an effect of Adcyap1 on longer fall migration duration (Mueller et al., 2011; Peterson et al., 2013). The strongest relationships between migratory traits and Adcyap1 were with an individual’s minimum allele length (Appendix 3). This may suggest dominance of the shortest Adcyap1 allele, a finding not reported in other species. However, the one individual in our sample that was homozygous for the longest Adcyap1 allele (163 bp) had by far the earliest spring departure, the longest spring and fall duration, and the latest fall arrival. This may support that the effects of long Adcyap1 alleles on migratory behavior are additive instead of dominant, as was previously suggested by Bourret and Garant (2015). Future studies with larger sample sizes may be able to more directly assess dominance by comparing migratory traits across genotypes at the individual level (Liedvogel et al., 2009; Saino et al., 2015). Our results that a single individual with a rare genotype differs greatly in migratory behavior are also consistent with studies published on other species that show a few individuals with rare genotypes can show significantly different migratory traits compared to the rest of the population (Bazzi et al., 2015). For example, a single individual barn swallow with the longest observed Clock allele had significantly later migration in both spring and fall compared to the rest of the population (Bazzi et al., 2015).

We found evidence of a significant interaction between Clock and Adcyap1 allele lengths on migratory duration, the first such finding in studies of migratory birds to our knowledge. Previous studies that have investigated the effects of both Clock and Adcyap1 either did not find a significant interaction between genes or did not test for one (Bazzi et al., 2016, 2017; Bourret & Garant, 2015; Contina et al., 2018; Peterson et al., 2013; Saino et al., 2015). In this interaction in our study, Adcyap1 length appears to increase migration duration when corresponding Clock allele length is short, especially in the fall. The peptide product of Adcyap1, PACAP, regulates the expression of Clock in chicken (Gallus gallus domesticus) pineal glands (Nagy & Csernus, 2007), which may suggest the interaction effect we observe is the result of differential expression of Clock as determined by Adcyap1 allele length. However, it is yet unknown how Adcyap1 allele length influences PACAP structure or function, including downstream regulation of Clock. These complex genetic interactions may explain in some cases why previous studies have failed to identify effects in single candidate genes.

Our findings that candidate gene allele lengths are correlated with five of the six migratory traits raise an important question about whether the migratory phenotypic characters assessed in this study and others are evolutionarily independent of one another. For example, is variation in arrival dates a secondary consequence of natural selection on departure date and migratory duration, or might selection be acting on these characters separately and independently? Among the 17 blackpoll warblers used in this study, duration and departure date for both spring and fall migration were strongly and negatively correlated (spring r = −0.948, fall r = −0.677), suggesting that birds who departed later had a shorter duration (faster rate of migration). However, duration and arrival date were not correlated or only moderately so for spring and fall, respectively (spring r = −0.010, fall r = 0.457). While a study with similar sample sizes in another species showed a close relationship between departure and arrival dates (Ouwehand & Both, 2017), surprisingly, departure and arrival dates were weakly correlated in our data for both seasons (spring r = 0.270, fall r = 0.290). These data suggest migratory departure and arrival dates may be independent of one another. Further, variation in duration may be more closely tied to variation in
whether these genes influence spring and fall duration via a common mechanism, perhaps through increased migratory restlessness (Mueller et al., 2011; Peterson et al., 2013), or whether these characters are independent requires further investigation. Nilsson et al. (2013) examined timing and speed of fall and spring migration in published tracking studies and found evidence for stronger selection on spring migratory phenology compared to fall, potentially suggesting evolutionary independence of these phenotypic characters.

Part of the answer to this question about evolutionary independence depends on the control of these characters at a proximate molecular level. Are there separate molecular pathways that control departure dates and duration, or that control spring and fall duration, or are these all proximately linked? PACAP is known to have broad influence on the physiology and behavior of organisms, acting in the brain and throughout peripheral organs (Mueller et al., 2011; Vaudry et al., 2009). It is therefore plausible that this gene could influence similar behaviors in separate seasons via independent molecular pathways. Conversely, it is perhaps equally plausible that a common molecular pathway is triggered by environmental cues in multiple seasons. This again highlights the need for studies that further investigate the functional role of candidate genes, how they influence migratory behaviors throughout the annual cycle, both independently and interactively with other factors (e.g., age, sex, environmental conditions), and specifically how variation in allele length influences the expression level, structure, and molecule functioning of gene products.

Although we found relationships between candidate genes and migratory traits at the individual level, we found no geographic structure in candidate genes. One possible explanation for this is that geographic variation in migratory behavior is explained by variation in environmental cues and not local adaptation in candidate genes. Geographic variation in behavior may be the result of plastic responses to variable environments, independent of the effects of those environments on selection in candidate genes (Foster, 1999). Therefore, to the extent environmental cues vary across breeding (or wintering) sites, it is possible we might observe behavioral differences in migratory behaviors without underlying geographic differences in genes. For example, during fall migration western breeding blackpoll warblers depart earlier and take longer to arrive at wintering grounds compared to eastern breeding birds (DeLuca et al., 2015, 2019), despite no differences in candidate gene frequencies between those populations. Our results suggest these differences across populations could be the results of plastic responses to differing environmental cues, while candidate genes are still correlated with variation in timing among individuals within each population.

**TABLE 4** General linear model results for individual-level analyses of blackpoll warbler candidate gene alleles and migratory traits

<table>
<thead>
<tr>
<th>Gene</th>
<th>Migratory trait</th>
<th>Allele</th>
<th>n</th>
<th>$\beta_{\text{Allele Length}}$</th>
<th>SE</th>
<th>$p_{\text{Allele Length}}$</th>
<th>$R^2_{\text{Allele Length}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clock</td>
<td>Spring arrival</td>
<td>Max</td>
<td>11</td>
<td>1.250</td>
<td>0.438</td>
<td>0.024</td>
<td>0.538</td>
</tr>
<tr>
<td></td>
<td>Spring departure</td>
<td>Min</td>
<td>12</td>
<td>5.615</td>
<td>1.577</td>
<td>0.007</td>
<td>0.618</td>
</tr>
<tr>
<td>Adcyap1</td>
<td></td>
<td>Min</td>
<td>12</td>
<td>5.123</td>
<td>1.657</td>
<td>0.015$^*$</td>
<td>0.544</td>
</tr>
<tr>
<td></td>
<td>Spring duration</td>
<td>Min</td>
<td>17</td>
<td>1.941</td>
<td>0.588</td>
<td>0.006$^*$</td>
<td>0.456</td>
</tr>
<tr>
<td></td>
<td>Fall duration</td>
<td>Mean</td>
<td>17</td>
<td>1.844</td>
<td>0.797</td>
<td>0.038$^*$</td>
<td>0.292</td>
</tr>
<tr>
<td></td>
<td>Fall arrival</td>
<td>Min</td>
<td>17</td>
<td>2.122</td>
<td>0.945</td>
<td>0.043</td>
<td>0.279</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>17</td>
<td>2.889</td>
<td>1.053</td>
<td>0.017</td>
<td>0.367</td>
</tr>
</tbody>
</table>

Note: Allele indicates whether an individual’s minimum, maximum, or mean allele length was used for each analysis. Slope ($\beta$), p-value, and partial $R^2$ values are given for allele length and each migratory trait. Population was also used as a predictor variable, though the slopes and p-values for this factor are not provided. Only models for which allele length was significant ($p \leq 0.05$) are shown. Results for all general linear models provided in Appendix 3.

*Significant single gene effects that are not further discussed because of a significant genetic interaction effect on this migratory trait (see Table 5).

**FIGURE 4** Relationship between individual maximum Clock allele length and spring arrival date residuals after accounting for population. Line represents slope from the general linear models with population and maximum Clock allele as independent variables. Individuals with longer maximum Clock allele lengths arrived earlier for their population compared to individuals with shorter maximum Clock alleles.

departure date than arrival date, perhaps due to stronger selection on arrival date especially in the spring (Nilsson et al., 2013). Our data also showed a strong positive correlation between spring and fall duration ($r = 0.688$), and both of these characters were associated with a significant interaction between Clock and Adcyap1 (Figure 6). Whether these genes influence spring and fall duration via a common mechanism, perhaps through increased migratory restlessness (Mueller et al., 2011; Peterson et al., 2013), or whether these characters are independent requires further investigation.
Further, Bourret and Garant (2015) point out that gene-by-environment interactions are underappreciated in the study of candidate genes. In their study of breeding phenology in tree swallows (Tachycineta bicolor), they found most candidate gene genotype-phenotype relationships were affected by environmental variables such as breeding density, latitude (a proxy of photoperiod), and temperature (Bourret & Garant, 2015). If genotype-phenotype relationships are influenced by environment, again we may observe behavioral differences across populations without underlying population genetic differences. Alternatively, it might simply be that any interpopulation variation in allele frequencies or lengths are undetectable given high levels of intrapopulation variation and our relatively small sample size; however, our result of no geographic structure of candidate genes is consistent with studies in other bird species (Bazzi et al., 2016; Dor et al., 2012, 2011; Johnsen et al., 2007). It is unlikely that the lack of geographic structure in candidate genes in blackpoll warblers is due solely to admixture from gene flow between locally adapted populations. Ralston and Kirchman (2012) found genetic structure due to isolation by distance at the continental scale in this species, despite a relatively high number of shared alleles among populations (Ralston & Kirchman, 2012, 2013). Future studies may benefit from comparing geographic patterns in candidate and neutral loci (Contina et al., 2018).

We observed greater allelic diversity in blackpoll warblers at Adcyap1 than at Clock, a pattern that appears to be typical across birds (Bazzi et al., 2016; Contina et al., 2018; Peterson et al., 2013; Johnsen et al., 2007). We present Table 5 and Figure 5 to illustrate these patterns for spring and fall duration.

**TABLE 5** General linear model results for interaction effects between candidate genes on spring and fall duration

<table>
<thead>
<tr>
<th>Clock allele</th>
<th>Adcyap1 allele</th>
<th>Migratory trait</th>
<th>( \beta_{\text{Clock} \times \text{Adcyap1}} )</th>
<th>( p_{\text{Clock} \times \text{Adcyap1}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>Min</td>
<td>Spring duration</td>
<td>−1.910</td>
<td>0.019</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fall duration</td>
<td>−0.740</td>
<td>0.007</td>
</tr>
<tr>
<td>Mean</td>
<td>Min</td>
<td>Spring duration</td>
<td>−0.143</td>
<td>0.934</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fall duration</td>
<td>−0.720</td>
<td>0.029</td>
</tr>
<tr>
<td>Max</td>
<td>Min</td>
<td>Spring duration</td>
<td>−1.450</td>
<td>0.056</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fall duration</td>
<td>−0.605</td>
<td>0.042</td>
</tr>
<tr>
<td>Mean</td>
<td>Spring duration</td>
<td>0.658</td>
<td>0.671</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Fall duration</td>
<td>−0.558</td>
<td>0.08</td>
<td></td>
</tr>
</tbody>
</table>

Note: Clock and Adcyap1 allele columns indicate whether an individual’s minimum, maximum, or mean allele length was used for each analysis. Slope (\( \beta \)) and p-value are given for the interaction term. Population was also used as a predictor variable, though the slope and p-values for this factor are not provided. GLMs were only run using mean or max Clock allele lengths and minimum or mean Adcyap1 allele lengths based on results from single gene models. Only models with spring or fall duration as the dependent variable are presented here; no other migratory traits had a significant interaction term. Models with significant interaction effects (\( p \leq 0.05 \)) are bolded.
Steinmeyer et al., 2009). However, our results differ in that we find relatively high heterozygosity (74.3%) at the Clock locus, despite only observing four alleles. In studies of other bird species, Clock heterozygosity was quite low (0%–9%; Bazzi et al., 2015, 2016; Chakarov et al., 2013; Dor et al., 2011). Low heterozygosity at microsatellite loci may be the result of stabilizing selection or loss of diversity due to inbreeding and therefore useful in studies of individual fitness (Chapman, Nakagawa, Coltman, Slate, & Sheldon, 2009; Dor et al., 2011). While a single microsatellite locus is not sufficient in estimating genome-wide heterozygosity or inbreeding (Miller et al., 2014), individual heterozygosity at a small number of microsatellite loci can be useful in the study of avian life history traits and fitness (Forstmeier, Schielzeth, Mueller, Ellegren, & Kempenaers, 2012; Lens et al., 2000), especially if those loci are functionally important and under selection (Chapman et al., 2009). Adcyap1 heterozygosity in Eurasian blackcaps was significantly associated with earlier spring arrival (Mettler et al., 2015), perhaps suggesting these individuals were of greater migratory fitness and able to migrate more quickly. We found similar results in blackpoll warblers in that heterozygotes at Adcyap1 had shorter spring and fall duration than homozygotes. We found no associations between heterozygosity at Clock and any migratory trait. Few studies of migratory candidate genes have examined the effect of individual heterozygosity, but our results and those of Mettler et al. (2015) suggest this is a potentially fruitful avenue of further investigation.

The results of our individual-level analyses demonstrate the value of studies that reveal individual migratory phenotypes, for example, by using light-level geolocators (McKinnon & Love, 2018). The combination of geolocator data with candidate gene analysis has been an important advancement in the study of migration (Bazzi et al., 2015; Contina et al., 2018; Saino et al., 2017). For example, in the study of barn swallows, individual-level analyses revealed that rare Clock genotypes can have a significant impact on migratory phenology (Bazzi et al., 2015) and that degree of DNA methylation at the Clock gene can explain individual variation in migratory behavior (Saino et al., 2017). Both of these insights would likely have been missed with analysis of population-level genetic and migratory variation. Increasing the number of species and individuals with linked genotype-phenotype information will allow richer investigations of migratory candidate genes.

Understanding variation in candidate genes across species is especially important for behaviors that are evolutionarily labile and may have arisen independently in multiple lineages. We therefore encourage future studies of candidate migratory genes to investigate species that are codistributed, share a biogeographic history, or are in the same family as those species that have already been studied. This will allow a better understanding of the influences of environment and history on selection at candidate genes, as well as the degree to which these patterns are conserved within lineages.

ACKNOWLEDGMENTS
We thank Christian Artuso, Yousif Attia, John Brett, Jukka Jantunen, Jill Boelsma, George Gress, David Merz, Jason Reppert, Emily Williams, and Ted Murphy-Kelly for outstanding assistance in the field. We are also grateful to the following organizations for support of this research: Saint Mary’s College, Kenneth M. Molson Foundation, Canadian Foundation for Innovation, Natural Sciences and Engineering Research Council of Canada, Wildlife Conservation
CONFLICT OF INTEREST
None declared.

AUTHOR CONTRIBUTIONS
JR, LL, and MM conceived of the project, performed the genetics laboratory work and statistical analyses, and wrote the paper. All other authors contributed to field efforts to collect blood samples or retrieve geolocators, and provided critical feedback on analyses and the manuscript. WVD took the lead on geolocator data analysis.

DATA ACCESSIBILITY
All data used in analyses in the paper, including Clock and Adcyap1 genotype information, capture coordinates, and migratory trait values are available in Dryad Digital Repository (https://doi.org/10.5061/dryad.d10qb58). An R script is available in supplementary materials (Data S1) and can be used in combination with the dataset on Dryad to perform all analyses in this paper.

ORCID
Joel Ralston https://orcid.org/0000-0002-7240-3322
Bradley K. Woodworth https://orcid.org/0000-0002-4528-8250

REFERENCES
**APPENDIX 1**

Tissue and blood sampling locations, population assignments for the current study, and original tissue sources (population, sample location, source, n)

<table>
<thead>
<tr>
<th>Population (n)</th>
<th>Source</th>
<th>Source n</th>
<th>Catalog numbers</th>
<th>Geocator numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alaska (14)</td>
<td>Current study (blood)(^a)</td>
<td>10</td>
<td>3254-001, 3254-003, 3254-008, 3254-011, 3254-005</td>
<td></td>
</tr>
<tr>
<td></td>
<td>University of Alaska Museum (frozen tissue)</td>
<td>3</td>
<td>UAM7394, UAM20089, UAM20508</td>
<td></td>
</tr>
<tr>
<td></td>
<td>UCLA Conservation Genetics Resource Center (feather)</td>
<td>1</td>
<td>01N9262</td>
<td></td>
</tr>
<tr>
<td>Yukon (11)</td>
<td>Current study (blood)(^a)</td>
<td>11</td>
<td>blpw12, blpw14, blpw15, blpw25</td>
<td></td>
</tr>
<tr>
<td>Manitoba (18)</td>
<td>Current study (blood)(^a)</td>
<td>17</td>
<td>4105-002, 4105-004, 4105-006, 4105-008, 4105-009, 4105-010, 4105-016, 4105-017</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ralston and Kirchman (2012; blood)(^b)</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eastern (29)</td>
<td>Current study (blood)(^a)</td>
<td>10</td>
<td>New York State Museum (frozen tissue)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ralston and Kirchman (2012; blood)(^b)</td>
<td>15</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^a\)Blood samples and DNA extracts archived at Saint Mary’s College, Notre Dame, Indiana, USA.

\(^b\)Blood samples and DNA extracts archived at New York State Museum, Albany, New York, USA.

**APPENDIX 2**

General linear model results for variation in allele lengths by longitude and latitude. There was no significant relationship between allele length for either locus and either longitude or latitude, regardless of whether individual minimum, maximum, or mean allele lengths were considered

<table>
<thead>
<tr>
<th>Gene</th>
<th>Allele</th>
<th>(\beta_{\text{Longitude}})</th>
<th>(p_{\text{Longitude}})</th>
<th>(\beta_{\text{Latitude}})</th>
<th>(p_{\text{Latitude}})</th>
<th>F</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clock</td>
<td>Min</td>
<td>0.005</td>
<td>0.720</td>
<td>0.004</td>
<td>0.544</td>
<td>0.220</td>
<td>2.69</td>
<td>0.804</td>
</tr>
<tr>
<td>Clock</td>
<td>Mean</td>
<td>0.008</td>
<td>0.466</td>
<td>0.018</td>
<td>0.759</td>
<td>0.417</td>
<td>2.69</td>
<td>0.664</td>
</tr>
<tr>
<td>Clock</td>
<td>Max</td>
<td>0.011</td>
<td>0.388</td>
<td>-0.006</td>
<td>0.926</td>
<td>1.437</td>
<td>2.69</td>
<td>0.245</td>
</tr>
<tr>
<td>Adcyap1</td>
<td>Min</td>
<td>-0.004</td>
<td>0.776</td>
<td>-0.055</td>
<td>0.417</td>
<td>0.579</td>
<td>2.69</td>
<td>0.563</td>
</tr>
<tr>
<td>Adcyap1</td>
<td>Mean</td>
<td>0.007</td>
<td>0.555</td>
<td>0.003</td>
<td>0.957</td>
<td>0.483</td>
<td>2.69</td>
<td>0.619</td>
</tr>
<tr>
<td>Adcyap1</td>
<td>Max</td>
<td>0.017</td>
<td>0.209</td>
<td>0.062</td>
<td>0.389</td>
<td>0.859</td>
<td>2.69</td>
<td>0.428</td>
</tr>
</tbody>
</table>
### APPENDIX 3

General linear model results for individual-level analyses of blackpoll warbler candidate gene alleles and migratory traits. Allele indicates whether an individual’s minimum, maximum, or mean allele length was used for each analysis. Slope ($\beta$) with standard error (SE), $p$-value, and partial $R^2$ are given for the relationship between allele length and each migratory trait. Population was also used as a predictor variable, though the slope and $p$-values for this factor are not provided. Models with a significant slope for allele length ($p \leq 0.05$) are shown with bold text.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Migratory trait</th>
<th>Allele</th>
<th>$n$</th>
<th>$\beta_{\text{AlleleLength}}$</th>
<th>SE</th>
<th>$p_{\text{AlleleLength}}$</th>
<th>$R^2_{\text{AlleleLength}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Clock</strong></td>
<td>Spring departure</td>
<td>Min</td>
<td>12</td>
<td>−0.258</td>
<td>1.811</td>
<td>0.890</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>12</td>
<td>−0.488</td>
<td>2.156</td>
<td>0.827</td>
<td>0.006</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Max</td>
<td>12</td>
<td>−0.810</td>
<td>2.478</td>
<td>0.752</td>
<td>0.013</td>
</tr>
<tr>
<td></td>
<td>Spring duration</td>
<td>Min</td>
<td>12</td>
<td>−0.333</td>
<td>1.752</td>
<td>0.854</td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>12</td>
<td>−0.353</td>
<td>2.09</td>
<td>0.870</td>
<td>0.004</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Max</td>
<td>12</td>
<td>−0.310</td>
<td>2.413</td>
<td>0.901</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>Spring arrival</td>
<td>Min</td>
<td>11</td>
<td>−0.626</td>
<td>0.384</td>
<td>0.147</td>
<td>0.275</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>11</td>
<td>−0.912</td>
<td>0.425</td>
<td>0.069</td>
<td>0.397</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Max</td>
<td>11</td>
<td>−1.250</td>
<td>0.438</td>
<td>0.024</td>
<td>0.538</td>
</tr>
<tr>
<td></td>
<td>Fall departure</td>
<td>Min</td>
<td>17</td>
<td>−0.102</td>
<td>0.543</td>
<td>0.854</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>17</td>
<td>−0.533</td>
<td>0.633</td>
<td>0.415</td>
<td>0.052</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Max</td>
<td>17</td>
<td>−0.923</td>
<td>0.599</td>
<td>0.147</td>
<td>0.154</td>
</tr>
<tr>
<td></td>
<td>Fall duration</td>
<td>Min</td>
<td>17</td>
<td>−0.369</td>
<td>0.577</td>
<td>0.534</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>17</td>
<td>0.065</td>
<td>0.702</td>
<td>0.927</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Max</td>
<td>17</td>
<td>0.662</td>
<td>0.679</td>
<td>0.347</td>
<td>0.068</td>
</tr>
<tr>
<td></td>
<td>Fall arrival</td>
<td>Min</td>
<td>17</td>
<td>−0.556</td>
<td>0.804</td>
<td>0.502</td>
<td>0.035</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>17</td>
<td>−0.574</td>
<td>0.967</td>
<td>0.563</td>
<td>0.026</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Max</td>
<td>17</td>
<td>−0.354</td>
<td>0.977</td>
<td>0.723</td>
<td>0.01</td>
</tr>
<tr>
<td><strong>Adcyap1</strong></td>
<td>Spring departure</td>
<td>Min</td>
<td>12</td>
<td>−5.615</td>
<td>1.577</td>
<td>0.007</td>
<td>0.613</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>12</td>
<td>−2.646</td>
<td>2.575</td>
<td>0.334</td>
<td>0.117</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Max</td>
<td>12</td>
<td>0.611</td>
<td>1.893</td>
<td>0.755</td>
<td>0.013</td>
</tr>
<tr>
<td></td>
<td>Spring duration</td>
<td>Min</td>
<td>12</td>
<td>5.123</td>
<td>1.657</td>
<td>0.015</td>
<td>0.544</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>12</td>
<td>2.435</td>
<td>2.509</td>
<td>0.360</td>
<td>0.105</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Max</td>
<td>12</td>
<td>−0.537</td>
<td>1.835</td>
<td>0.777</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td>Spring arrival</td>
<td>Min</td>
<td>11</td>
<td>−0.742</td>
<td>0.543</td>
<td>0.214</td>
<td>0.211</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>11</td>
<td>−0.322</td>
<td>0.645</td>
<td>0.633</td>
<td>0.034</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Max</td>
<td>11</td>
<td>0.106</td>
<td>0.462</td>
<td>0.826</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td>Fall departure</td>
<td>Min</td>
<td>17</td>
<td>0.122</td>
<td>0.738</td>
<td>0.871</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>17</td>
<td>0.522</td>
<td>0.866</td>
<td>0.557</td>
<td>0.027</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Max</td>
<td>17</td>
<td>0.470</td>
<td>0.631</td>
<td>0.470</td>
<td>0.041</td>
</tr>
<tr>
<td></td>
<td>Fall duration</td>
<td>Min</td>
<td>17</td>
<td>1.941</td>
<td>0.588</td>
<td>0.006</td>
<td>0.456</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>17</td>
<td>1.844</td>
<td>0.797</td>
<td>0.038</td>
<td>0.292</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Max</td>
<td>17</td>
<td>0.512</td>
<td>0.681</td>
<td>0.465</td>
<td>0.042</td>
</tr>
<tr>
<td></td>
<td>Fall arrival</td>
<td>Min</td>
<td>17</td>
<td>2.122</td>
<td>0.945</td>
<td>0.043</td>
<td>0.279</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>17</td>
<td>2.889</td>
<td>1.053</td>
<td>0.017</td>
<td>0.367</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Max</td>
<td>17</td>
<td>1.500</td>
<td>0.878</td>
<td>0.111</td>
<td>0.183</td>
</tr>
</tbody>
</table>