UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

CONSERVATION, ECOLOGY, AND EVOLUTION OF MIGRATORY BEHAVIOR IN THE PAINTED BUNTING (*Passerina ciris*)

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

ANDREA CONTINA
Norman, Oklahoma
2015

CONSERVATION, ECOLOGY, AND EVOLUTION OF MIGRATORY BEHAVIOR IN THE PAINTED BUNTING (*Passerina ciris*)

A DISSERTATION APPROVED FOR THE
DEPARTMENT OF BIOLOGY

BY

_____
Dr. Jeffrey Kelly, Chair

_____
Dr. Eli Bridge

_____
Dr. Bruce Hoagland

_____
Dr. Michael Patten

_____
Dr. Thomas Smith

_____
Dr. Larry Weider

Dedication

To anyone who wants to learn how this planet works.

Acknowledgements

My work at the University of Oklahoma and at the Oklahoma Biological Survey was supported by a National Science Foundation EAGER grant through my advisors Jeff Kelly and Eli Bridge from 2010-2012. Additional funding was provided by the Department of Biology through teaching assistantships, the George M. Sutton Scholarship in Ornithology, and the Robert E. and Mary B. Sturgis Scholarship of Merit.

The following wildlife managers provided field access at the Wichita Mountains Wildlife Refuge: U.S. Fish and Wildlife Service managers; Tony Both, Walter Munsterman, Dan McDonald, Scott Johnson, and Jeremy Dixon.

University of Oklahoma staff in the Biology Department and at the Oklahoma Biological Survey, particularly Trina Steil, Ranell Madding, Kaye Carter, Kyle Baker, Carol Baylor, Robbie Stinchcomb, Dan Hough, George Davis, and George Martin, provided logistic support and outstanding help with the necessary paperwork. Ann Harris greatly helped me in the Biology CORE Molecular Laboratory at OU. Members of the Kelly and Bridge lab provided helpful comments and encouragement over the years. Jeff Kelly, Eli Bridge, and Frank Moore (University of Sothern Mississippi) provided help with the needed permits and equipment for banding. Additional help and valuable suggestions have been provided by Rosemary Knapp, Rich Brougthon, Kristen Ruegg (UCSC), Ryan Harrygan (UCLA), John Pollinger (UCLA). My friends and colleagues at OU Claire Curry, Diane Landoll, Roeder Karl, David Hille, Kyle Horton, Kyle Broadfoot, Jonathan Kruk, James Cureton, Jelena Bujan, Billy Culver, Brent Tweedy, Rachel Hartnett, FeiFei Zhang, Dagmar Frisch, Phillip Morton.

My committee members have been always helpful with advice, constructive comments, and encouragement. I want to thank (in alphabetical order) Eli Bridge, Bruce Hoagland, Michael Patten, Thomas Smith, and Larry Weider for investing so much time in helping me to be a better scientist, and for reviewing multiple drafts of this dissertation. Jeff Kelly has been a fantastic advisor, I have learned a lot from him, and I will never thank him enough for his advice and guidance over the years.

Table of Contents

List of Tables

Chapter 1.

Chapter 2.

Chapter 4.

Chapter 5.

List of Figures

Chapter 1.

Chapter 4.

Chapter 5

Abstract

In animal migration studies, determining the migratory connectivity of different species is a crucial step to understand the ecology and evolution of complex behavioral traits as well as to implement effective conservation management plans of declining species. Intrinsic markers, such as stable isotopes and genetic markers, in conjunction with miniaturized light-level geolocator devices, provide high resolution tools to study the ecology and evolution of migration in birds smaller than 20 grams and offer insight on the breeding population of origin of those individuals trapped and sold in the international pet trade market. For example, from 2005 to 2009, CITES documented more than 317,000 live birds sold annually within an industry estimated to be worth billions of US dollars per year. In this context, migratory species can be particularly difficult to protect because any management effort would necessarily involve cooperation among different countries with distinct laws and regulations. A case in point is the Painted Bunting (*Passerina ciris*), which breeds primarily in the United States and winters in Mexico, Central America, southern Florida, and the Caribbean. I provide a new library of polymorphic microsatellite loci and offer a new tool for genetic population structure studies, forensic analyses, and conservation management of the Painted Bunting and other related species of the genus *Passerina*. My results suggest that buntings harvested in Central America for the retail market belong to the western breeding population as opposed to the smaller Atlantic breeding population. However, a percentage of both populations come into contact at the wintering grounds in the Yucatan Peninsula where conservation efforts might need to be strengthened.

Furthermore, during my investigations on the molting ecology of the Painted Buntings, I found that birds breeding in Oklahoma and molting in Sinaloa, Mexico, show a clear bimodal pattern of stable isotope ratios in their flight feathers, as previously demonstrated in this species, suggesting that some birds may initiate molt immediately upon arrival in northwestern Mexico whereas others may delay molt. I used geologger tags to test whether differences in the timing and route of fall migration movements were related to stable isotope signatures in primary feathers. I observed variation among individuals in migration routes, wherein birds from the same breeding population differed greatly in their use of molting and wintering locations. However, I did not find a relationship between isotope signatures and the timing or route of fall migration.

Furthermore, I investigated possible relationships between variation in two candidate genes implicated in the control of migratory behavior (CLOCK and ADCYAP1) and several aspects of fall migration including initiation date, arrival at molt/stopover sites, and duration of the first leg of fall migration. I evaluated the mutation rate of these candidate genes through an Approximate Bayesian Computation (ABC) coalescent approach and studied the effect of natural selection on their allele frequencies across populations. I found that millennial scale summer temperatures and longitude likely affected the current genotypes of CLOCK and ADCYAP1 in three sampling populations. To search for novel genes associated with migratory performance, I implemented 454 next generation sequencing and generated over 48,000 DNA sequences distributed over the Painted Bunting genome. I assembled an extensive library of candidate genes by annotations extracted from the Mouse Genome Informatics database and I focused on the following candidate genes: ADRA1d,

ANKRD17, CISH, and MYH7. I tested these genes across avian migratory and non-migratory species, revealing a surprising degree of allelic variation in some genes. MYH7 correlated with speed of migration (km/day) and body mass across 11 species of songbirds in addition to two novel uncharacterized polymorphic genes that strongly correlated with duration of migration in the Painted Buntings.

Chapter 1: Using geologgers to investigate bimodal isotope patterns in Painted Buntings (*Passerina ciris*)

This chapter is published, with some modifications, as Contina, A., E.S. Bridge, N. E. Seavy, J. M. Duckles, and J. F. Kelly. 2013. Using geologgers to investigate bimodal isotope patterns in Painted Buntings (*Passerina ciris*). The Auk 130(2): 265-272.

Abstract

Painted Buntings (*Passerina ciris*) that breed in Oklahoma and molt in Sinaloa, Mexico, demonstrate a clear bimodal pattern of stable isotope ratios in their flight feathers. Some birds had a C3 carbon signature in primary 1 (P1, the first feather replaced during wing molt) and a C4 carbon signature in primary 9 (P9, the last primary to molt), whereas other sympatric birds evinced a C4-based diet throughout feather molt. The bimodal pattern of stable isotope ratios in flight feathers suggests that some birds may initiate molt immediately upon arrival in northwestern Mexico (and carry a C3 signature with them from the breeding grounds) whereas others may delay molt (and grow feathers solely from C4 plants of Sinaloa). From 2010 to 2012, we used geologger tags to test whether differences in the timing and route of fall migration movements were related to stable isotope signatures in primary feathers. We analyzed stable isotopes of hydrogen and carbon in P1 and P9 from 25 individuals fitted with geologger tags in two consecutive years. Of these, 60% changed the diet (C3 vs. C4) that was used to grow P1 between years. We also observed variation among individuals in migration routes, wherein birds from the same breeding population differed greatly in their use of molting and wintering locations. However, we did not find a relationship between isotope

signatures and the timing or route of fall migration. We speculate that the bimodal isotope signature we observed represents a carryover effect related to local landscapes (grassland or agriculture vs. shrubland) used during the late breeding season and early molting period, and that these effects diminish as molt progresses. If this is the case, there is the potential for breeding-season diet to directly affect plumage quality in this molt migrant.

Introduction

One important element of migration and molt strategies is the ability of birds to track seasonal shifts in food availability (Stach et al. 2012, but for other taxa see also Bischof et al. 2012, and Fryxell and Avgar 2012). Focusing on a population of Painted Buntings breeding in Oklahoma, Bridge et al. (2011) found a bimodal distribution of carbon isotope ratios in primary flight feathers (P1 and P9) indicating a midmolt shift from a C3 diet (derived from C3 photosynthesis) to a C4 diet (derived from C4 photosynthesis). Such a pattern suggests that two groups of individuals in this population could have fundamentally different migratory routes, such that the two carbon signatures reflect two molting locations, one with C3 food resources and the other with C4 resources. To test this hypothesis, Bridge et al. (2011) examined carbon stable isotope ratios in newly grown feathers from birds in a single stopover molting area in Sinaloa, Mexico. Feathers at this location also had a bimodal carbon signature, which suggests that differences in molting locations are not necessary to explain the differences in diet (i.e., isotope ratios of feathers).

An alternative explanation for the bimodal pattern is that, depending on the timing and duration of migration, stable carbon isotope ratios of P1 can reflect the diet on the

breeding ground, which in Oklahoma is dominated by C3 plants, or the diet at the molting location, which is dominated by C4 plants. Although the mechanism behind this relationship could take several forms, the most parsimonious explanation would be that early migrants arrive at the stopover region early and spend a week or more equilibrating with the local trophic web prior to molt, whereas late migrants begin molt immediately after arriving in the molting area, such that they grow their first primary using reserves that reflect isotope ratios from the breeding grounds. In this case, we would predict that the carbon isotope signature of P1 would be associated with the date of arrival on the molting grounds.

Here, I provide the first quantitative descriptions of the migration biology of individual Painted Buntings using data from custom-designed, low-cost geologgers. This is the smallest North American migrant tracked to date and among the first molt-migrant species to be tracked using geologgers. In particular, I used light data collected from individual Painted Buntings to estimate (1) the date of initiation of migration, (2) the number of days between initiation of migration and stopover for molt, (3) the date of arrival at the molting grounds, and (4) routes of migratory movements. I used these data to test the prediction that C3 carbon signatures in P1 occur more often in late migrants that, presumably, molt immediately after arrival on the molting grounds and that C4 carbon signatures are common in early migrants that may remain on the molting grounds for several days or weeks prior to molting. In addition, I measured stable isotope ratios in these birds' feathers in two consecutive years so that I could examine year-to-year consistency of isotope ratios in feathers, which is presumably related to consistency in diet and molting locations. On the basis of known distributions of

Painted Buntings throughout the year (Lowther et al. 1999), I predicted that the migration route of birds from our study population in Oklahoma would involve a westward flight to northwestern Mexico to molt, followed by a southward movement and subsequent stationary period within the documented wintering range. If bimodal patterns in isotope values indeed arise from timing of molt, rather than differences in migratory pathways, I expected that all birds from our breeding population would follow similar migratory trajectories, but that the isotope ratios of the first primary would be related to the timing of migration.

Methods

Field site and tagging effort

I studied a population of Painted Buntings that bred at the Wichita Mountains Wildlife Refuge (N34.4–W98.4), Oklahoma. This refuge is largely dominated by mixed prairie grasses (mostly C4 species) such as Little Bluestem (*Schizachyrium scoparium*), Indian grass (*Sorghastrum nutans*), and Switchgrass (*Panicum virgatum*), but Painted Buntings are located primarily in patches of forest characterized by oaks (*Quercus* spp.) and Eastern Red cedar (*Juniperus virginiana*) that represent an important source of dietary C3 carbon (Eskew 1938, Martin et al. 1951, Askins 1993, Carter et al. 2008). I deployed 120 geologgers on free-ranging male Painted Buntings during summer 2010 and 2011 (May–July) and released them shortly after capture. We used solar-powered geologgers of our own design that weighed 0.6 to 0.7 g (including harness material), representing ~4% of the body mass of this species. The geologgers measured light intensity every minute and recorded the average light measurement every 10 min on an arbitrary scale from zero to 127. I mounted the tags using leg-loop harnesses (Rappole

and Tipton 1991) made of 0.7-mm elastic silicon beading thread (Stretch Magic brand; Pepperell Braiding, Pepperell, Massachusetts) that were presized (see Naef-Daenzer 2007) and assembled by heat welding (rather than tying) prior to deployment. I also tagged each bird with individually numbered federal bands and plastic color bands. Adult singing males were captured in 12-m mist nets close to nests or active territories using playback calls and decoy mounts.

To study the effect of migration timing on migration route, I deployed 80 additional geologgers (40 in 2010 and another 40 in 2011) on individuals captured in May of each year and assigned to a manipulative experiment, in which birds were held in captivity under different light regimes before being fitted and released with geologgers. I designed this manipulative experiment to accelerate or decelerate the annual cycle of the birds. Half of the birds belonged to an "early" group, which was released on 1 July of each year, and the other half belonged to a "late" group that was released on 1 August (Table XX). Recaptured birds with geologgers from year 1 were fitted with new geologgers in year 2 to enable us to document migratory behavior of single individuals in consecutive years. Thus, I deployed a total of 200 geologgers over the course of the study.

Stable isotopes: carbon and hydrogen

The stable isotope signatures of primary feathers reflect nutrient deposition into keratin tissue at the beginning and end of the molting period. Thus, I measured stable isotope ratios in the innermost primary (P1, first to be molted) and the outermost primary (P9, last to be molted) collected from the right wing upon each recapture event in summer 2011 and 2012. I analyzed the $^{13}C/^{12}C$ ratio ($\delta^{13}C$) and deuterium/hydrogen ratio (δD) at

the University of Oklahoma following the laboratory procedures detailed in previous publications (Kelly et al. 2009, Paritte and Kelly 2009, Bridge et al. 2011). I report our results in delta notation in relation to PeeDee Belemnite for $\delta^{13}C$ and in relation to Standard Mean Ocean Water for δD (Craig 1957, 1961). I used generalized linear models to test for correlation between variables and, in particular, the association between $\delta^{13}C$ signature in P1 molted in 2011 and three explanatory variables: (1) departure dates from the breeding ground in Oklahoma, (2) arrival dates at the molting sites in northwestern Mexico, and (3) duration of migration from Oklahoma to the molting grounds in days. All statistical analyses were performed using XLSTAT. I considered feathers with $\delta^{13}C$ values greater than –18‰ to be primarily C4 in origin, based on Bridge et al. (2011), and feathers with δD values greater than –65‰ to be of potential Mexican origin, based on our interpretation of δD distribution proposed by Hobson et al. (2009).

Light-level geolocator tags

To derive location estimates from raw light-level data, I employed the threshold method implemented in the R package GeoLight (R Development Core Team 2005, Lisovski and Hahn 2012). For purposes of illustration, I also conducted an exploratory analysis to estimate migratory routes using the curve-fitting method implemented in the R package tripEstimation (Sumner et al. 2009, Sumner and Wotherspoon 2010). Detection of twilight events is a crucial step in defining the correct location of birds with geologgers. With the threshold method, I determined twilight events for the GeoLight package using a threshold value of 6.5 on the scale of zero to 127, which corresponded to a sun angle of –3 to –4.5 (depending on the individual tag), based on an

individual calibration period that corresponded to a minimum of 3 days after deployment while the birds were still resighted on the breeding grounds. The first 24 h after deployment were regarded as a behavioral equilibration period, and I did not use these data for tracking or calibration. To compute migration departure and arrival dates, I then plotted the coordinates for each data point in ARCGIS and visualized the longitudinal and latitudinal movements associated with each day of the year. A consistent longitudinal movement (e.g., ~5 consecutive data points with longitudinal shift ≥5°W) was considered a real movement and not an artifact of shading. I did not consider latitude when computing departure and arrival dates because of the high degree of error in latitude estimates.

I implemented tripEstimation by following the general steps described in Seavy et al. (2012). I began by truncating the 0–127 raw light measurements to measurements from 6 to 20 that captured the twilight period. I discarded all twilight transitions with substantial shading, and designated tag-specific calibrations using the same periods from the threshold method. I constrained possible locations with a land mask, the known locations of release day, and the spatial boundaries beyond which I considered locations unrealistic (values outside of latitude 0 to 40 or longitude –120 to –40). I disregarded latitudinal estimates during the equinox period (15 days on either side of the true equinoxes). For the light parameters, I used variance in light data = 2.5, variance in light attenuation = 6, an Ekstrom range of –20 to 40, and variance outside this range = 7. For our movement model, I used a log normal distribution with a mean of 2.6 km h^{-1} and variance of 1.3. I set up the Markov-chain Monte Carlo (MCMC) to start by drawing 10,000 samples for burn-in and tuning of the proposal distribution. I

then repeated an MCMC from the end of the burn-in by drawing 10,000 samples and made sure that the model parameters were tuned in for each tag and that the MCMC had converged by assessing minimal migration trajectory shift from the burn-in period and the additional run of 10,000 simulations. Finally, I set up a new MCMC by drawing 10,000 simulations and then generating a last draw of 10,000 samples to describe the posterior distribution. I used the mean of the posterior distribution coordinates to plot our estimate of the most likely routes for each individual. Results are presented as means ± SD.

RESULTS

In 2011, I retrieved 13 of the 100 geologger tags deployed in 2010, but only 2 geologgers lasted to the following year of recapture to provide usable data (see Table xx). During summer 2012 I retrieved 32 of 100 geologger tags deployed in 2011. Of these, 24 provided usable data through fall migration (e.g., November).

Overall, of the 26 tags with usable data, 23 came from free-ranging birds tagged in 2011, 2 from birds used in my photoperiod manipulations trial carried out in 2010, and 1 from my photoperiod manipulations carried out in 2011. One individual was recaptured in consecutive years (2011, 2012).

Migratory geography

Of the 25 birds with geologger data, 23 migrated from the breeding ground in southwestern Oklahoma toward the molting site in northwestern Mexico (Fig. 1A and supplemental material). The “noisy” light measurements yielded only approximate estimates of geographic position because, as is typical of current technology, poor weather conditions (e.g., clouds), feather shading of the light sensor, and behavior all

influenced light levels recorded by the geologger. Nevertheless, the overall westward migratory pattern from Oklahoma toward Mexico was clear. There were two notable exceptions to this pattern: one unmanipulated bird migrated east to northern Louisiana (ID 251112862 in supplemental material, Fig. S1). His location after that was unknown because the geologger failed in mid-September. Another bird, a member of the "late" experimental group in 2010, migrated immediately south along the eastern edge of the Gulf of Mexico and spent the winter at or near the Isthmus of Tehuantepec before returning to the breeding ground in the spring along the same route (ID 222150719 in Fig. 1D). Some individuals appeared to stop over in southwest Texas (about a week, possibly along the Rio Grande) during their migration to the Sinaloa–Sonora region (see supplemental material).

Most geologgers stopped collecting data while birds were still in the Sinaloa–Sonora region. However, I recovered breeding-season-to-breeding-season tracks from 2 individuals used in light manipulation trials in 2010 and a track until late January from 1 wild bird tagged in 2011. As described above, 1 bird from the "late" light manipulation treatment migrated immediately south to wintering grounds in southern Mexico and then returned to Oklahoma in spring (ID 222150719 in Fig. 1D). The other bird from the "late" light manipulation treatment and the unmanipulated bird both followed similar loop migration paths, beginning with southwestward movement to northwest Mexico (Sonora and Sinaloa regions) in late summer, followed by less rapid southward movement in late fall, and an eastward shift in early winter to southeastern Mexico (ID 222150714 in supplemental material and ID 222150755 in Fig. 1B, C, and table xx, respectively). There were also variations in movements associated with later

stages of their migration. In particular, the bird from the 2010 light manipulation group showed an eastward movement to the Isthmus of Tehuantepec (or thereabout) and a direct northward spring migration in early May to return to the breeding site, whereas the wild individual tagged in 2011 showed a longer eastward movement and wintered primarily in the Yucatan Peninsula.

Migratory timing

Painted Buntings departed the Oklahoma breeding site from mid-July through mid-August (Table S1). In 2010, one bird migrated to the molting site in 4 days (the other migrated eastward). In 2011, the duration of migration ranged from a few days to as long as 29 days and averaged 11.4 days (Tables S1 and S2).

Migration timing and diet shifts

The bimodal pattern in carbon isotope ratios in the first primary feathers previously observed by Bridge et al. (2011) was evident in samples from 2011 and 2012 (Figs. 2 and 3). I found no significant association between $\delta^{13}C$ signatures in primary feathers (P1) molted in 2011 and timing of migration variables when considered together in a single model or independently. The full model describing variation in $\delta^{13}C$ signature in P1 molted in 2011 as a function of departure dates from the breeding ground in Oklahoma, arrival dates at the molting sites in northwestern Mexico, and duration of this migration in days was not significant ($R^2 = 0.14$, ANOVA; $F = 0.98$, df = 3 and 17, $P = 0.421$). Similarly, $\delta^{13}C$ values in P9 molted in 2011 had no relationship with departure dates from Oklahoma, arrival dates at the molting sites in Mexico, or duration of fall migration ($R^2 = 0.16$, ANOVA; $F = 1.12$, df = 3 and 19, $P = 0.316$). I also tested each of these relationships in a univariate regression, and none was significant ($P =$

0.714, n = 22). The results suggest that the bimodal pattern in carbon stable isotope in P1 is not associated with departure dates from the breeding ground or pace of migration. Interestingly, comparisons of feathers taken from the same bird in consecutive years revealed that 60% of the birds switched "pattern" between years as reflected in the carbon isotope ratios of P1. For P1s molted in 2010, 50% of the birds had a carbon signature compatible with a diet based on C3 carbon sources typical of mesic plants (i.e., < $\delta^{13}C$ = –18‰), whereas only 15% of the birds had a carbon signature in P9 compatible with a diet based on primarily C3 food intake. The mean $\delta^{13}C$ value for P1 molted in 2010 was –18.3 ± 4.1‰ (range: –23.3 to –12.0‰), and that for P9 molted in 2010 was –14.2 ± 2.9‰ (range: –22.1 to –10.5‰). However, my examination of $\delta^{13}C$ signature in primary feathers molted in the same birds in 2011 showed that 80% of the birds had a carbon signature in P1 compatible with a diet based on C3 carbon sources, and only 5% had a carbon signature in P9 compatible with a diet based on C3 carbon sources. The mean $\delta^{13}C$ value for P1 molted in 2011 was –19.1 ± 5.1‰ (range: –22.6 to –10.8‰), and that for P9 molted in 2011 was 13.9 ± 3.9‰ (range: –23.0 to –11.4‰).

DISCUSSION

Molting schedules can differ between populations of the same or closely related species. Examples include eastern and western populations of the Painted Bunting, along with eastern and western populations of Warbling Vireos (*Vireo gilvus*) and Northern Rough-winged Swallows (*Stelgidopteryx serripennis*) (Thompson 1991a, Yuri and Rohwer 1997, Voelker and Rohwer 1998). There is also evidence that migration and molt can vary among individuals within populations (Elrod et al. 2011). For the western Painted Bunting, bimodal isotopic signatures in a sample of primary feathers

from a single breeding population provided highly suggestive evidence for the existence of variation in migration behavior among individuals (Bridge et al. 2011). This variation compels testing of hypotheses regarding the causes of multiple migratory patterns and/or diverging migratory schedules within a breeding population of Painted Buntings in Oklahoma.

My investigation confirmed that the frequency distribution of carbon isotope ratios in the first primary was bimodal in Painted Buntings that molted at stopover locations in west Mexico, which indicates that some birds used primarily C3 food and others used C4 sources early in molt. This bimodal pattern contrasts with the frequency distribution of carbon isotope values in the ninth primary, which was uniformly C4 in origin. I tested whether the bimodal distribution in P1 was the result of variation in (1) migration routes or (2) migration timing, and found no support for either hypothesis. Hypotheses that remain to be tested include the possibility that the carbon isotope ratio of P1 reflects the composition of the breeding territory in Oklahoma or of habitats used in the early stages of migration, reinforcing the importance of the role played by molting behaviors and migratory patterns in processes leading to speciation (Rohwer and Irwin 2011).

The across-year consistency in $\delta^{13}C$ is interesting, considering that so many birds evidently altered their use of C3 and C4 food sources between years, and suggests that local effects, as opposed to landscape effects, led to the bimodal distribution of $\delta^{13}C$. That is, the landscape was consistent between years, giving rise to similar between-year averages, but the use of particular patches within the landscape (perhaps agricultural fields vs. grasslands) may have given rise to bimodal $\delta^{13}C$ population

distributions. I note that in my research I focused on adult males to maximize the rate of recaptures. However, we should consider the possibility that migratory patterns of females and juveniles may differ greatly, as previously demonstrated in other species (Butler et al. 2002, Jenkins and Cristol 2002, Bai and Schmidt 2012).

Although there was temporal variation in the migratory behavior of Painted Buntings in my study population, it was not related to the bimodal distribution of $\delta^{13}C$ in primary feathers. My data also demonstrate large year-to-year variation in $\delta^{13}C$ and δD in isotope ratios of primary feathers of individuals. Thus, it is evident that the different isotope profiles for P1 and P9 do not constitute genetically "hard-wired" migration or dietary strategies. I found no clear relationship between isotope signatures and migration timing. Moreover, the frequency with which birds evinced changes in dietary carbon sources between years indicates that rigid dietary preferences were not likely a cause of the bimodal distribution of isotope signatures in P1 or the transition to unimodal pattern by the end of molt.

I also documented individual variation in departure dates from the breeding ground, arrival dates in northwestern Mexico, the time used to migrate between these two sites, and the duration of short refueling stops in southwestern Texas. However, it is unclear why two adult birds of the same sex and from the same breeding population diverged so dramatically from the general migratory directional pattern shown by the other birds. Perhaps environmental instability related to the monsoons of the Sonora–Sinaloa region (Rohwer et al. 2005, Pyle et al. 2009) has promoted the evolution and maintenance of a degree of flexibility in the migratory behavior of Painted Buntings and other avian migratory species that breed in the Great Plains of the United States (Jahn et

al. 2013). Also, the severe drought in southwestern Oklahoma in July 2011 (National Oceanic and Atmospheric Administration) likely had an effect on timing of migration in Painted Buntings and many other species of grassland birds.

The power of combining light-based geolocation data and other markers such as stable isotopes is becoming increasingly evident in investigations of migratory species (González-Solís et al. 2011). Carbon stable isotopes are widely used to infer the spatial distribution of consumers and in migratory connectivity studies (Marra et al. 1998, Hobson et al. 2012). My findings suggest that the carbon isotope signatures of primary feathers are related to diet late in the breeding season or early in migration, independent of location. This inference was possible only because I incorporated information from the geologgers. Indeed, my threshold approach provided a relatively easy procedure to extrapolate variables associated with the timing and duration of migration, but it is coarse in its resolution of movements and accuracy of locations. By contrast, my preliminary analysis of migration routes through the curve-fitting method (Fig. 1C, D) is an example of how we can estimate movements between simple data points. However, it can be computationally intensive and requires careful selection of model parameters.

In conclusion, my data suggest a marked convergence to a similar use of resources toward the end of the molting period, as evinced by stable isotope analysis of P9 feathers, and that the bimodal carbon signature in P1 (Bridge et al. 2011) is not related to variation in arrival date at the molting locations and subsequent timing of the onset of molt. The most likely remaining explanation is that P1 is grown using energy reserves acquired prior to migration. This is an intriguing possibility because this would

be a quantitative and direct carryover effect from the breeding season to the migration and molting stages of the Painted Bunting's life history. Such a carryover might have ecological and evolutionary implications (Marra and Holberton 1998, Norris and Marra 2007). For example, the degree to which plumage quality is a useful proxy in selection of breeding partners may depend on the information it contains about the ability of individuals to garner resources on the breeding ground in the previous year rather than at stopover sites encountered during migration (Keyser and Hill 2000).

Acknowledgments

I thank my field assistants C. Randall, A. Stark, K. May, S. Collins, H. Stuart, T. Michels, and C. Griffin; the managers of the Wichita Mountains Wildlife Refuge, in particular T. Booth, W. Munsterman, and J. Dixon, for their collaboration and logistic support; and Alex Jahn and Michael Murphy for comments that greatly improved the manuscript. This research was funded by the National Science Foundation (awards DEB 0946685 and DBI 1152356 to E.S.B. and J.F.K.). Manipulation of captive birds was in accordance with the Institutional Animal Care and Use Committee of the University of Oklahoma. Data on the July 2011 drought are available from the National Oceanic and Atmospheric Administration, National Climatic Data Center, at www.ncdc.noaa.gov/sotc/drought/2011/7.

Figures and tables

Table S1. The category "Start DOY" indicates the beginning of migration from the breeding ground (Wichita Mountains Wildlife Refuge, Oklahoma), expressed as day of year, of individual Painted Buntings (n = 26; † = tagged in 2010, ‡ = recaptured and retagged in 2011; ASY = after-second-year; M = male sexed by plumage).

Bird	Band ID[a]	Start migration	Start DOY[b]	Age	Sex
PABU-01	222150758	17-Jul-2011	198	ASY	M
PABU-02	222150789	19-Jul-2011	200	ASY	M
PABU-03	222150751	19-Jul-2011	200	AHY	M
PABU-04	251112864	21-Jul-2011	202	ASY	M
PABU-05	251112866	22-Jul-2011	203	ASY	M
PABU-06	222150770	22-Jul-2011	203	ASY	M
PABU-07	222150755	25-Jul-2011	206	ASY	M
PABU-08	222150752	26-Jul-2011	207	ASY	M
PABU-09	222150754	26-Jul-2011	207	ASY	M
PABU-10	251112762	26-Jul-2011	207	ASY	M
PABU-11	251112861	27-Jul-2011	208	ASY	M
PABU-12	251112838	28-Jul-2011	209	ASY	M
PABU-13	251112859	29-Jul-2011	210	ASY	M
PABU-14	251112883	29-Jul-2011	210	ASY	M
PABU-15	222150760	31-Jul-2011	212	ASY	M
PABU-16	251112865	31-Jul-2011	212	ASY	M
PABU-17	251112858	31-Jul-2011	212	ASY	M
PABU-18	251112854	31-Jul-2011	212	ASY	M
PABU-19	251112852	31-Jul-2011	212	ASY	M
PABU-20	251112819	2-Aug-2011	214	ASY	M
PABU-21	251112890	6-Aug-2011	218	ASY	M
PABU-22	251112862	8-Aug-2011	220	ASY	M
PABU-23	222150757	14-Aug-2011	226	ASY	M
PABU-24 †	222150714	23-Aug-2010	235	ASY	M
PABU-25 ‡	222150714	25-Aug-2011	237	ASY	M
PABU-26	222150719	1-Sep-2010	244	ASY	M

[a] U.S. Fish and Wildlife Service aluminum band.

[b] Mean ± SD = 209.7 ± 8.4.

Table S2. The category "Arrival DOY" indicates the arrival date in northwestern (NW) Mexico, expressed as day of year of individual Painted Buntings (n = 26; † = tagged in 2010; ‡ = recaptured and retagged in 2011).

Bird	Arrival in NW Mexico [a]	Arrival DOY [b]	Latitude of arrival in NW Mexico [c]	Longitude of arrival in NW Mexico [d]	Migration days [e]	Migratory remarks [f]
PABU-01	19-Jul-2011	200	31.58	–109.66	03	Bird used in Fig. 1A
PABU-02	31-Jul-2011	212	34.45	–108.13	13	Early group, 2011
PABU-03	3-Aug-2011	215	35.29	–111.68	15	
PABU-04	19-Aug-2011	231	24.25	–112.30	29	
PABU-05	31-Jul-2011	212	34.45	–108.13	10	
PABU-06	4-Aug-2011	216	28.10	–113.21	14	Bird used in Fig. 1A
PABU-07	3-Aug-2011	215	32.17	–110.69	10	Bird 1 in Fig. 1B, C
PABU-08	2-Aug-2011	214	18.70	–108.16	08	
PABU-09	9-Aug-2011	221	23.87	–109.34	15	Bird used in Fig. 1A
PABU-10	10-Aug-2011	222	19.88	–104.36	16	Bird used in Fig. 1A
PABU-11	6-Aug-2011	218	24.24	–105.50	11	Bird used in Fig. 1A
PABU-12	7-Aug-2011	219	28.78	–108.29	11	Bird used in Fig. 1A
PABU-13	31-Jul-2011	212	–1.39	–106.39	03	Bird used in Fig. 1A
PABU-14	7-Aug-2011	219	20.10	–113.54	10	Bird used in Fig. 1A
PABU-15	5-Aug-2011	217	26.13	–106.72	06	Bird used in Fig. 1A
PABU-16	5-Aug-2011	217	22.76	–110.97	06	
PABU-17	7-Aug-2011	219	17.54	–111.03	08	Bird used in Fig. 1A
PABU-18	13-Aug-2011	225	28.30	–111.02	14	Bird used in Fig. 1A
PABU-19	16-Aug-2011	228	23.33	–108.67	17	
PABU-20	23-Aug-2011	235	11.45	–111.80	22	Late group, 2011
PABU-21	11-Aug-2011	223	0.65	–104.69	06	
PABU-22	13-Aug-2011	225	34.65	–93.64	06	East migration, Louisiana
PABU-23	15-Aug-2011	227	15.86	–107.29	02	
PABU-24†	26-Aug-2010	238	35.08	–109.30	04	Late group, 2010
PABU-25‡	30-Aug-2011	242	23.19	–108.79	06	Late group, 2011
PABU-26	n/a	n/a	n/a	n/a	n/a	Bird 2 in Fig. 1D. Late group, 2010

[a] Sonora and Sinaloa.
[b] Mean ± SD = 219.9 ± 8.6.
[c] First day in NW Mexico. Latitude was not used in calculation of arrival dates in Mexico because of poor accuracy.
[d] First day in NW Mexico.
[e] Mean ±SD = 11.4 ±6.4.
[f] Light regime group.

Figure 1. Painted Bunting migration revealed by geologgers. Breeding and wintering grounds are highlighted in green and yellow, respectively. (A) Unfiltered data points for 12 birds are superimposed to show a clear westward movement and a shift in longitude (~10°) occurring at the end of July 2011. (B) Unfiltered data points for one bird migrating to the Yucatan peninsula in January 2012. (C) Migratory route for the same bird shown in panel B, but modeled for only the first part of the migration until August 2011. (D) Migratory route for one bird that migrated south, overwintered in Mexico, and returned to Oklahoma in spring 2012.

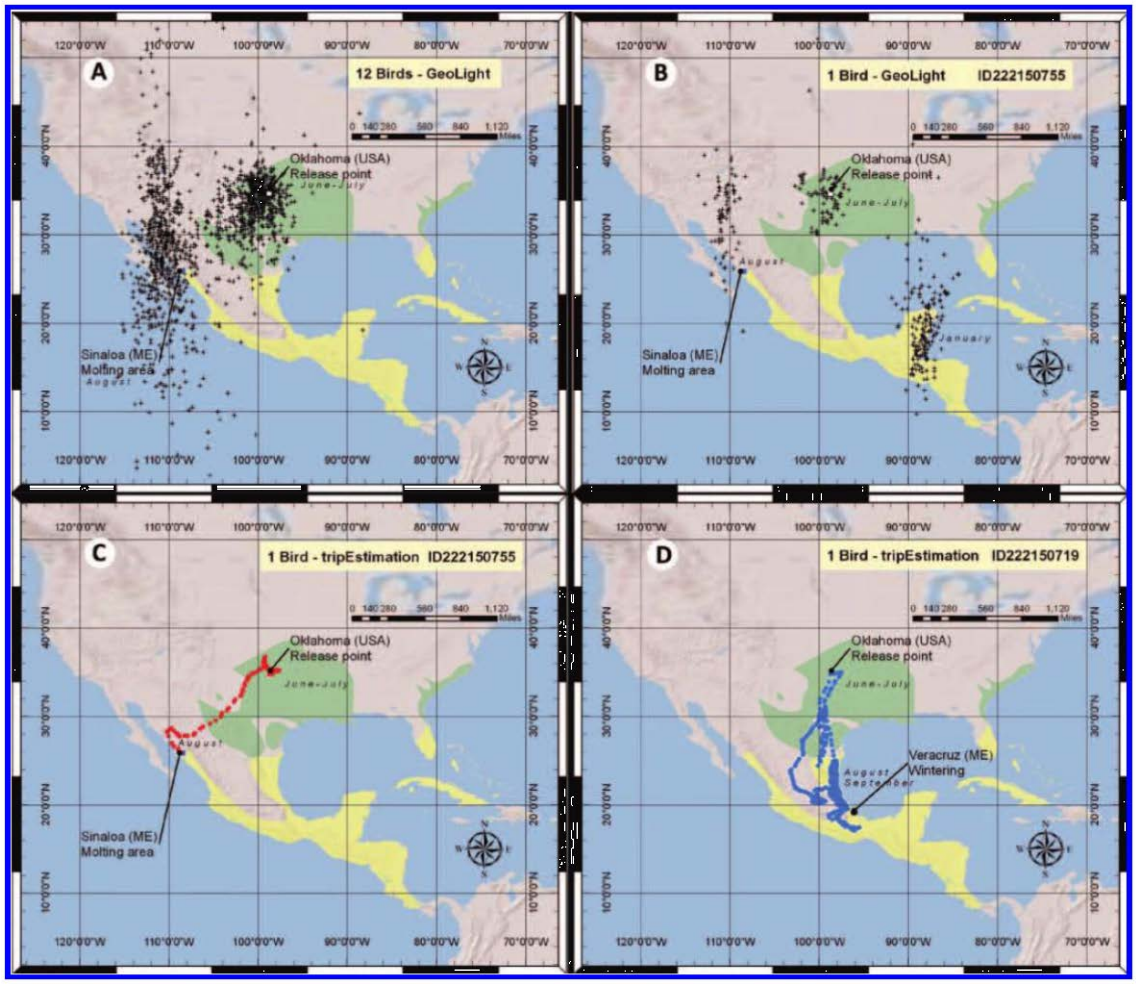


Figure 2. Distribution of (A, B) carbon and (C, D) hydrogen stable isotope ratios in primary feathers for individual Painted Buntings sampled in two consecutive years at the breeding site in Oklahoma (Wichita Mountains Wildlife Refuge). Feathers collected in summer 2011 (white bars) were molted in 2010, and feathers collected in 2012 (black bars) were molted in 2011.

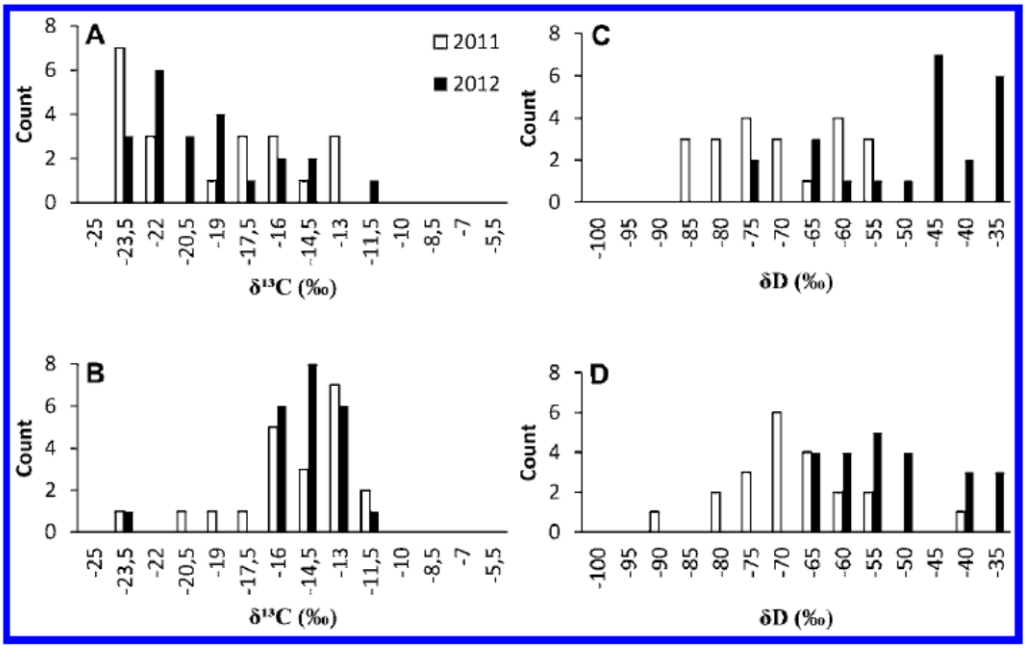


Figure 3. Histograms indicating carbon stable isotope ratios in P1 associated with the length of migration and departure dates from the breeding ground of Painted Buntings at Wichita Mountains Wildlife Refuge in Oklahoma (OK). Black bars represent free-ranging birds, white bars represent carbon stable isotope ratios in P1, and green bar indicates east migration (Louisiana); the photoperiod manipulation late group (migration onset delayed) is represented by red bars, and the photoperiod-manipulation early group and south migration (migration onset accelerated) are represented by blue bar. For a description of photoperiod manipulations (delayed and accelerated treatments), see details in the text.

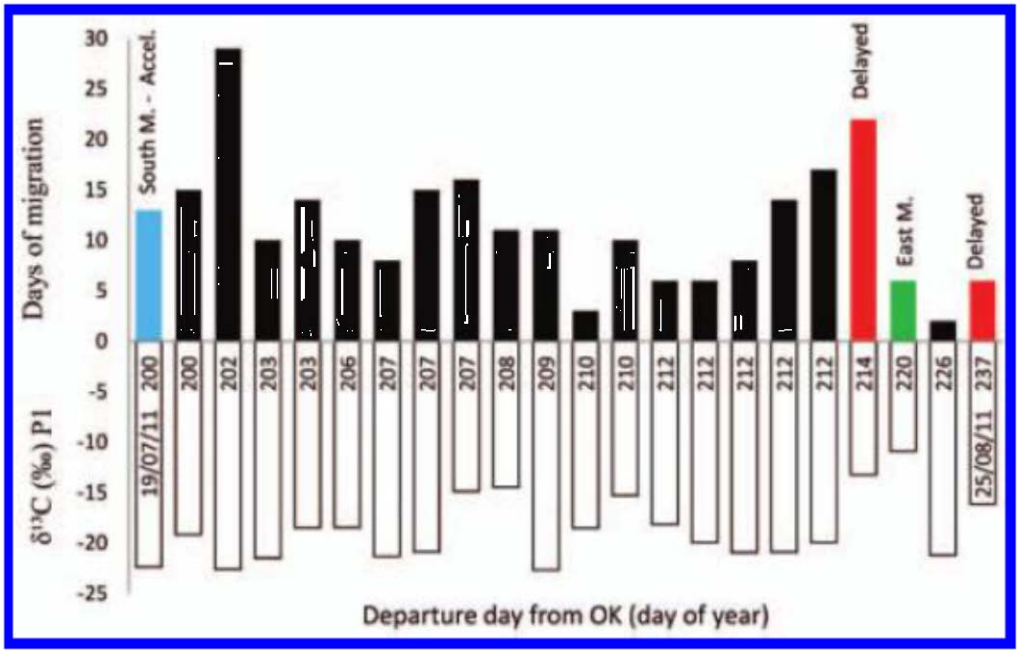


Figure S1. Painted Bunting migration revealed by geologgers: threshold-method results (n = 26; GeoLight output). Breeding and wintering grounds are represented by green and yellow shades, respectively. The label on the top right corner of each map indicates the individual U.S. Fish and Wildlife Service band number. Results shown are for the 2011–2012 migration, except for the last two maps (2010–2011 migration).

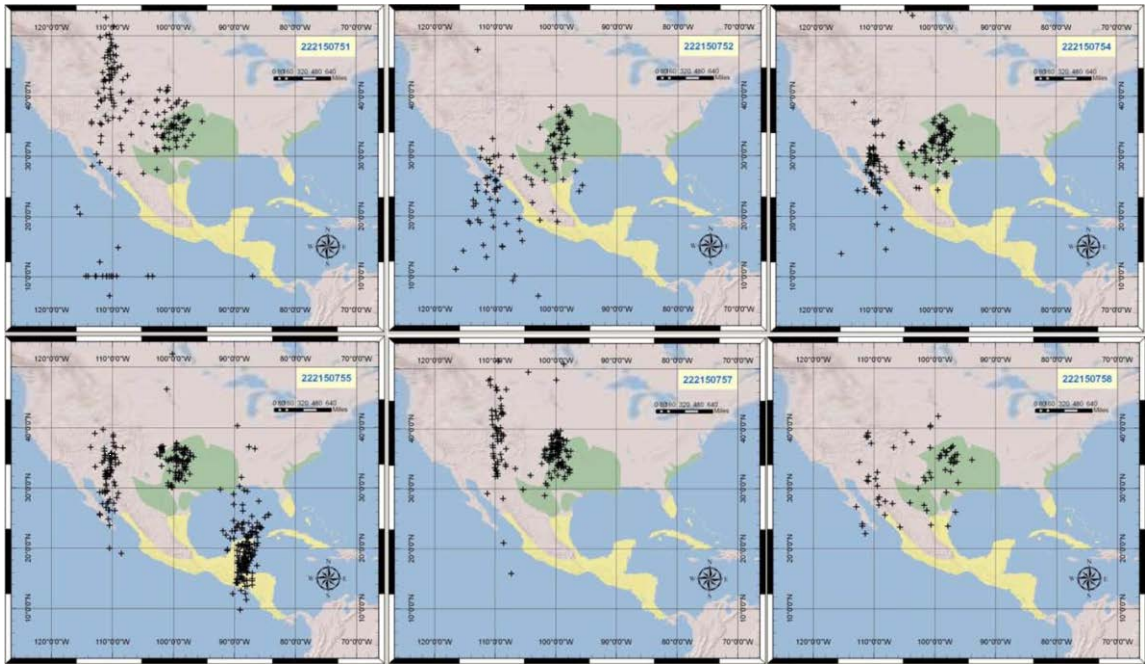


Figure S1. Continued

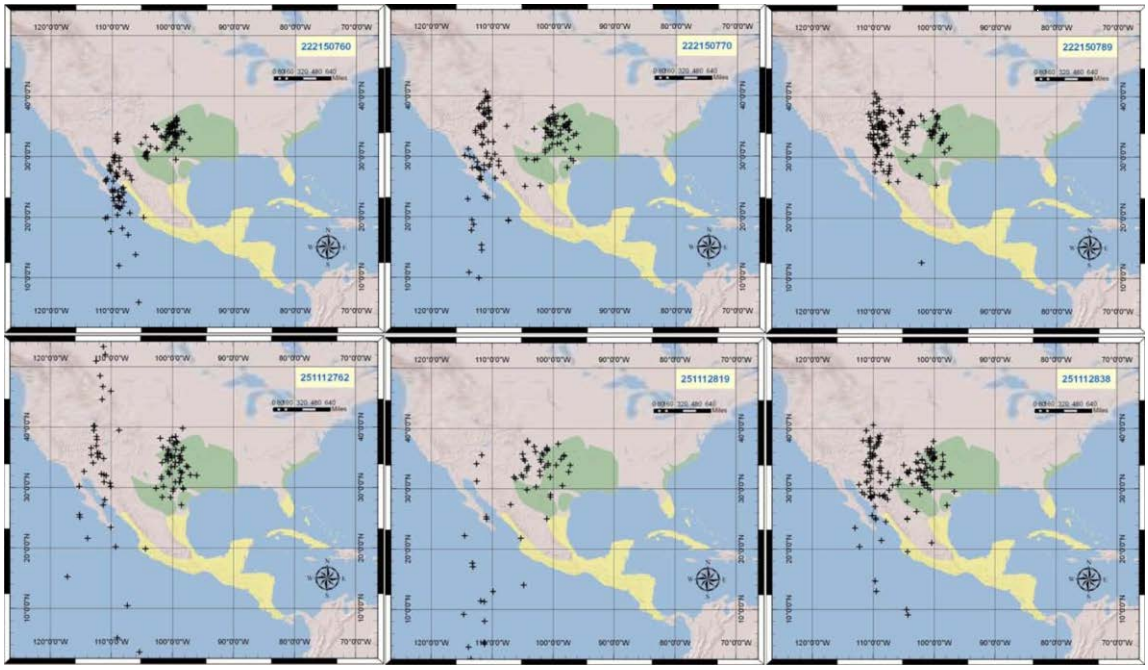


Figure S1. Continued

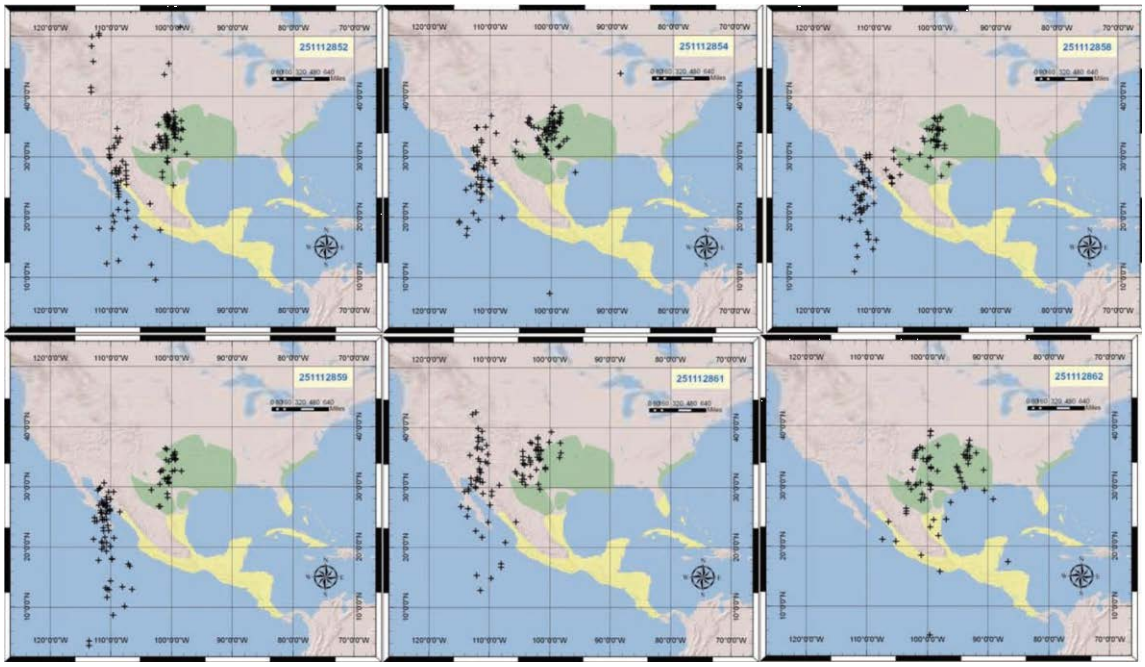
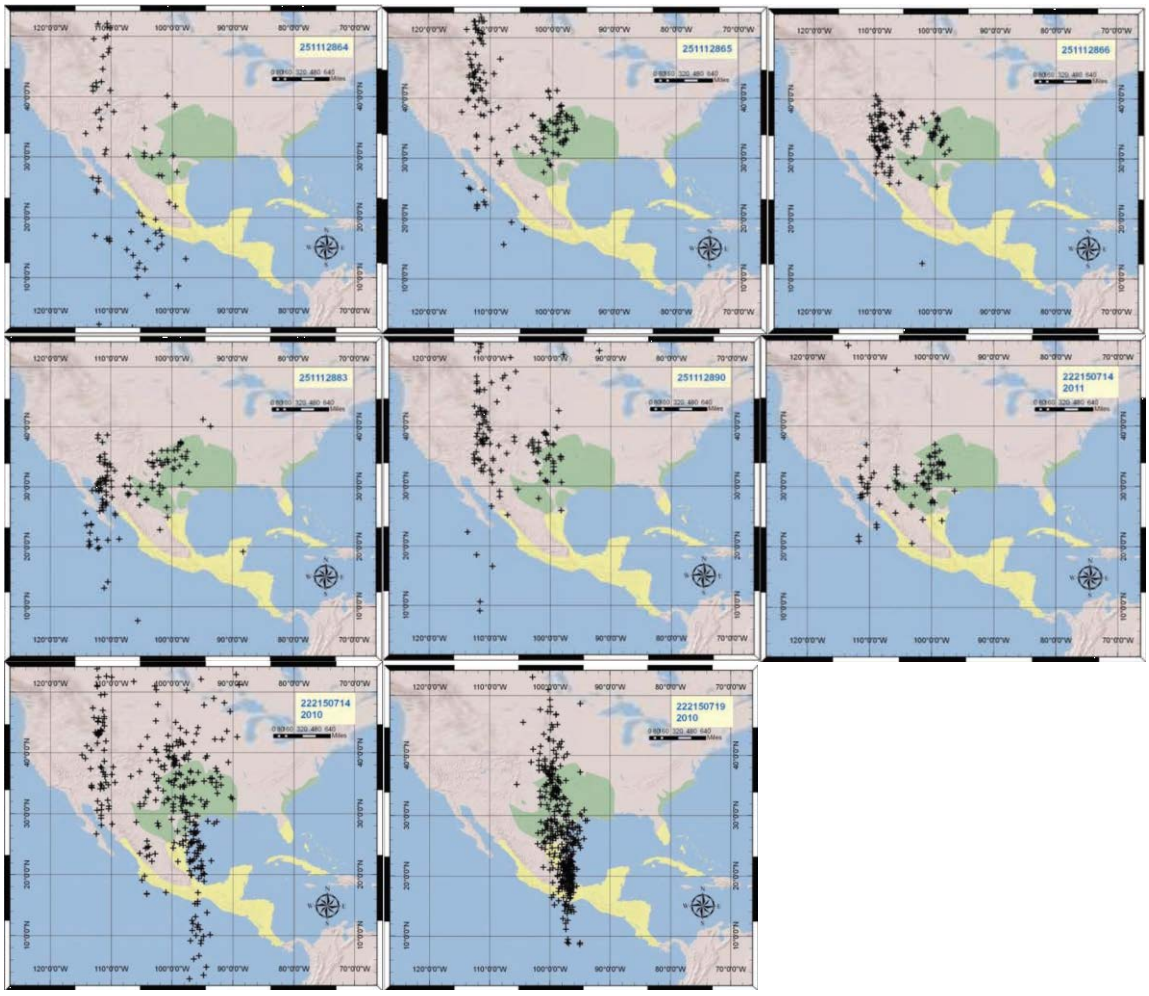


Figure S1. Continued



Chapter 2: Characterization of 15 microsatellites isolated in the Painted Bunting (*Passerina ciris*) for population and forensic analyses

---

This chapter is being submitted, with some modifications, to the journal Molecular Ecology Notes.

Abstract

I describe fifteen microsatellite loci (tetra-nucleotides) isolated from *Passerina ciris*, which is a Neotropical migratory bird distributed in the United States and Central America and is steeply declining in some parts of its range. I detected two to thirty two alleles per locus and the expected heterozygosity ranged from 0.21 to 0.85. I obtained positive results with cross-amplification in two other closely related species including *Passerina versicolor* and *Cardinalis cardinalis* (the Varied Bunting and the Northern Cardinal, respectively). I provide a new library of polymorphic microsatellite loci and offer a new tool for genetic population structure studies, forensic analyses, and conservation management of the Painted Bunting and other related species of the genus *Passerina.*

Rationale of the study

Passerina buntings is a genus (family Cardinalidae) that includes 6 species of small passerines, mostly Neotropical migrants (Sibley and Monroe 1990). *Passerina ciris* is a colorful species with an interesting geographical distribution that consists of two disjunct populations; a larger breeding population in the southwestern part of the US and northeastern part of Mexico, and a smaller breeding population narrowed distributed along the Atlantic coast of the US from North Carolina to Florida (Thompson 1991a;1991b). Previous genetic studies tried to determine the level of

differentiation among geographically distant populations of Painted Buntings by using mtDNA and calling for further analysis incorporating genetic markers of higher resolution such as DNA microsatellites and SNPs (Herr et al. 2011). Furthermore, at least over the last three decades, this species suffered a steep and constant decline in some part of its distribution range and it is thought to be heavily impacted by illegal pet trade (Iñigo-Elias et al. 2002). Here I present a novel library of 15 microsatellite loci isolated in the Painted Bunting aiming to provide a new set of high resolution genetic markers for further research on this species ranging from genetic population studies to forensic analysis of samples from the international illegal pet trade.

Methods

I extracted and purified genomic DNA from blood and muscles using QIAGEN DNA extraction kit according to the protocols supplied by the manufacturer whereas DNA from feathers was purified following a modified feather digestion procedure (30h at 60 °C) and DNA extraction mix with no DDT (modified from De Volo et al. 2008). I used a 454 high-throughput sequencing approach to generate 48 694 short sequences and found 185 variable regions presenting tetra-nucleotide repeats. I designed 40 microsatellite candidate primers using PRIMER3 software (http://frodo.wi.mit.edu/) (Rozen & Skaletsky 2000). I followed a hybrid primer technique to reduce the costs of molecular laboratory work (Boutin-Ganache et al. 2001). I implemented a M13 hybrid primer process with a combination of three primers: a hybrid forward primer with 16 bp of M13F-20 sequence (GTA AAA CGA CGG CCA G) tacked on the 5’ end, the 16bp M13F dye-labeled primer for which we used the 6FAM fluorescent dye (Applied Biosystems, CA) and the reverse primer. I performed PCR reactions in Eppendorf

Mastercycler under the following conditions: total volume 10ql, initial denaturation step of 95 °C for 15 min followed by 25 cycles of 94 °C for 30 s; 58-59 °C for 90 s; and 72 °C for 60 s. These first set of cycles were followed by 20 additional cycles of 94 °C for 30 s; 53 °C for 90 s; and 72 °C for 60 s. We included a final extension of 30 min at 60 °C to ensure the addition of a terminal adenine to all the PCR products (Brownstein et al. 1996). I analyzed amplicons using an ABI 3730 sequencer with an internal size standard (Genescan LIZ-600, Applied Biosystems, CA) and called allele sizes after electropherograms were visualized on PEAKSCANNER 2.0 software (Applied Biosystems, CA). I found fifteen polymorphic alleles out of the forty screened and initially tested on a few individual. Then I genotyped a total of 67 Painted Bunting samples collected across three geographically distant sites: Sinaloa (N = 15; Mexico, N25.1 – W107.3), Oklahoma (N = 43; Wichita Wildlife National Refuge – USA, N34.4 – W98.4), and Georgia (N = 9; N31.5 - W81.3). I also genotyped 9 and 2 samples of the closely related species Varied Bunting and Northern Cardinal, collected in New Mexico and Oklahoma (N = 9; San Andres National Wildlife Refuge – USA, N32.4 - W106.3, and N = 2; Wichita Wildlife National Refuge – USA, N34.4 – W98.4), respectively. I carried out Hardy–Weinberg equilibrium exact test and linkage disequilibrium tests using Genepop 3.4 (Raymond & Rousset 1995). Values of significance were estimated by the Markov chain method with 10 000 batches. I detected two to thirty-two alleles per locus were and the expected heterozygosity ranged from 0.21 to 0.85. All loci conformed to the Hardy–Weinberg expectations with no heterozygosis excess (Table 1). The Pb6 locus presented heterozygote deficit, probably indicating the occurrence of null alleles, non-random sampling, or inbreeding. Two pairwise tests of genotypic

disequilibrium (Pb2×Pb17, Pb6 ×Pb31) were significant (*P* value 0.035 and 0.001, respectively) and may indicate that these loci could be linked. I successfully cross-amplified 15 and 7 loci in the Varied Bunting and in the Northern Cardinal, respectively. My research offer a novel set of primers to further investigate genetic population structure of *Passerina ciris* and potentially other closely related species of this genus.

Acknowledgments

I thank Ann Harris for her valuable support in the Biology Core Molecular Lab at the University of Oklahoma. Hira Walker and The New Mexico Department of Game and Fish - Wildlife Management Division for providing Varied Bunting samples. Sievert Rohwer and Chris Wood for providing Painted Bunting samples from Sinaloa, Mexico. We thank our field assistants C. Randall, A. Stark, K. May, S. Collins, H. Stuart, T. Michels, and C. Griffin; The Wichita Mountains Wildlife Refuge; T. Booth, W. Munsterman, and J. Dixon, for their collaboration and logistic support.

Table 1

Characterization of fifteen microsatellite loci isolated from *Passerina ciris*. Number of alleles (na), observed (HO) and expected (HE) heterozygosity based on a sample of 67 individuals.

					PABU Sinaloa pop. n = 15				PABU Oklahoma pop. n = 43				PABU Georgia pop. n = 9			
Locus		Primer sequence (5′–3′)	Size (bp)	Repeat motif	Alleles n.	Ho	He	P value	Alleles n.	Ho	He	P value	Alleles n.	Ho	He	P value
Pb1	F:	ATCGATGCAATGGCTGCAC	242	(CCCT)^5	10	0.86667	0.78391	0.2913	9	0.67442	0.66594	0.7482	4	0.66667	0.66013	0.3555
	R:	CGTGCCTCAGTTTCCCATC														
Pb2	F:	ACCAGGTGATTCTGGTCTTC	235	(GGTT)^5	7	0.8	0.8069	0.3479	11	0.7907	0.76799	0.7871	6	0.77778	0.84314	0.5704
	R:	TGTTGCTTCAGCCCTTTGG														
Pb3	F:	TGAAGTACACATCTGACTCATGG	325	(AAAC)^5	16	0.73333	0.93563	0.9924	23	0.95349	0.94911	0.4672	8	0.55556	0.88235	0.9945
	R:	GAGCTCTAAAGTCATGATTAATGCAAC														
Pb4	F:	GGATAGGGTCTTCAAGCACAG	151	(AGAT)^5	7	0.6	0.70805	0.7672	14	0.81395	0.80958	0.6433	3	0.11111	0.45098	1
	R:	TGCCTGCTTATGTTATGGGC														
Pb5	F:	TGAGCATGGATCTGCCTAC	252	(GGAT)^5	14	0.93333	0.85287	0.2554	20	0.88372	0.84979	0.852	6	0.66667	0.77778	0.9406
	R:	GGAGAGGAGGGACTTGTGG														
Pb6	F:	TCAGAATTACCTGGTGCTCCC	207	(GAGT)^6	10	0.4	0.84138	0.9999	14	0.62791	0.82517	0.9997	5	0.33333	0.75163	0.9885
	R:	ACAGAGCCATTGGTCCAGC														
Pb7	F:	GGCTGAGTTTGGGCTATGG	256	(CTTT)^5	13	0.73333	0.77701	0.7506	19	0.69767	0.81642	1	5	0.88889	0.79085	0.2766
	R:	ACATGTGAGACATCACTAAAGGAG														
Pb8	F:	CAGCCCAAATGTTGGTTTGAC	219	(AAAT)^6	5	0.4	0.45287	0.9795	9	0.4186	0.50862	0.9865	4	0.88889	0.67974	0.1317
	R:	GCACACCAATGGGAGCATAC														
Pb9	F:	GTGTTGAGAGCCTGATGCC	161	(AATG)^14	15	0.73333	0.92644	0.9963	21	0.76744	0.91874	0.9998	12	0.77778	0.93464	0.9673
	R:	TCCAGGAGTGTGCCTAACATC														
Pb10	F:	GCAGTTCAGTTCCATCTGGC	238	(AAAT)^6	7	0.46667	0.62989	0.9945	9	0.39535	0.56689	0.9886	4	0.33333	0.70588	0.9971
	R:	TGCAACGTGGCTTTATTTGC														
Pb11	F:	GGGCTGAGACCTGCTTGG	193	(GGAT)^6	17	1	0.92644	0.2812	32	0.95349	0.95513	0.4226	9	0.88889	0.92157	0.6932
	R:	CCCAAACCATTGCCTGTGC														
Pb12	F:	CACAGCTCGGTGTCATTGG	214	(AATC)^7	10	0.53333	0.88276	0.9981	7	0.53488	0.7078	0.998	4	0.55556	0.71242	0.9247
	R:	ACAAACTCCTTAATAGCTGAGGC														
Pb13	F:	CCGCAATAAGGAGGTGGAAAC	190	(CCTT)^4	10	0.66667	0.84828	0.9129	15	0.74419	0.81778	0.8994	6	0.66667	0.7451	0.8216
	R:	ACCTCTGAGCAATTTCTGTAGC														
Pb14	F:	ATATTCCATGTTCTGATTAGTACCC	165	(ACAT)^4	3	0.6	0.51111	0.6667	3	0.66667	0.48046	0.1024	2	0.5	0.5	n/a
	R:	ACTGAGAAACCTGGTCTAGTGG														n/a
Pb15	F:	AGGATTGATGCGTTAGGCTTG	207	(AAAT)^5	3	1	0.64444	0.127	4	1	0.71667	0.0573	2	1	1	n/a
	R:	GGTCCTGGGCTCACCTTC														n/a
Pb16	F:	TCACAATGGCACACGAAGC	275	(GTTT)^6	8	1	0.95556	0.7932	13	0.875	0.975	0.9861	2	1	1	n/a
	R:	TGCAGACTTCAGGGACGTG														n/a
Pb17	F:	GCACAGGTCCATACTGAGAAC	177	(CCTT)^4	2	1	0.6	0.4001	2	1	0.6	0.4	n/a	n/a	n/a	n/a
	R:	TTTCTTTGCAGATGTGAAGGG														

Chapter 3: The international trade of Neotropical migratory birds; a genetic insight on the Painted Bunting case.

This chapter is being submitted, with some modifications, to the journal Conservation Biology.

Abstract

The Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES) is a worldwide effort among governments to monitor and safeguard international trade of wild animals and plants with the ultimate goal of preventing their extinction. However, from 2005 to 2009, CITES documented more than 317,000 live birds sold annually within an industry estimated to be worth billions of US dollars per year. In this context, migratory species can be particularly difficult to protect because any management effort would necessarily involve cooperation among different countries with distinct laws and regulations. A case in point is the Painted Bunting *(Passerina ciris*), which breeds primarily in the United States and winters in Mexico, Central America, southern Florida, and the Caribbean. Male Painted Buntings are brightly colored, which makes them highly sought after as pets, particularly in Mexico, Central America, and Europe. Although linking migrant songbirds legally or illegally captured outside the US boundaries to breeding populations is often a challenge that remains unresolved, there are genetic tools that can support these investigations. Here I describe our use of an extensive microsatellite DNA library to determine the population of origin of Painted Buntings sold in European and Mexican pet markets. My results suggest that buntings harvested in Central America for the retail market belong to the western breeding population as opposed to the smaller Atlantic breeding population.

However, a percentage of both populations come into contact at the wintering grounds in the Yucatan Peninsula where conservation efforts might need to be strengthened. I hope that my research and results will draw attention to the need for international collaboration and cooperation in combating illegal pet-trade activity.

Introduction

Species Background and Distribution

The Painted Bunting (*Passerina ciris*) is a songbird that winters in Mexico, Central America, and the Caribbean area (including the southern Florida), and migrates to the United States to breed (Rappole and Warner 1980; Lowther et al. 1999). The species' range encompasses the jurisdiction of 11 countries (Fig 1), all of which are currently partners of the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES). Two large populations reproduce in the US and northeast Mexico over the summer, but the breeding distribution is allopatric, and the larger western population, does not come into contact with the smaller and steeply-declining eastern population (Fig. 1; Thompson 1991a; Thompson 1991b; NAS 2004b). On the basis of measured traits and plumage coloration, Storer (1951) claimed two subspecies met along a suture zone running through the middle of the western range (dashed line in Fig. 1), giving an expanded range to "*P. c. ciris*" of the East and ignoring the species' discontinuous geographical distribution. By contrast, Herr et al. (2011) carried out a genetic study of the Painted Bunting in its breeding range employing mtDNA and found no evidence for the presence of a suture zone for two subspecies within the boundaries of the western species distribution, but supported the

notion that a declining population along the Atlantic coast should be considered an evolutionarily significant unit (ESU) for more efficient management of the species. Furthermore, the presence of a 500km wide gap that separates the two populations in their US breeding range is rare among species of North American songbirds (Fig 1) although species with disjunct breeding areas are found in populations of the Burrowing owl (*Athene cunicularia*) and Crested Caracara (*Caracara cheriway*), for example. Using niche modeling, Shipley et al. (2013) concluded that several important biotic and abiotic parameters could be compatible with a continuous breeding distribution, as seen in the case of the closely-related Indigo Bunting (*Passerina cyanea*) species. Their results suggested that segregated Pleistocene climatic refugia may still be reflected in the current distribution of the Painted Bunting. However, despite this distribution gap, it remains unresolved whether or not the two allopatric populations occupy the same overwintering sites and what routes they follow during spring and fall migration. Therefore, determining whether or not population overlap occurs during the wintering months is important for the implementation of more effective conservation strategies.

Population Decline and Conservation Concerns

The Painted Bunting has declined over the last five decades, in particular, in the eastern part of its breeding range and is considered as a species of conservation concern (Sauer et al. 1997; Sykes and Holzman 2005; Sykes et al. 2007; USFWS 2008). However, this decline is not homogeneous and some populations in the western and southeastern range are experiencing a demographic expansion (see Sauer et al. 2003 for a detailed population trend map). Currently, the causes of population declines are thought to be habitat loss (both at wintering and stopover sites), habitat fragmentation

(breeding grounds), Brown-headed Cowbird (*Molothrus ater*) nest parasitism, and illegal pet trade involving capture of large numbers of birds at their wintering grounds (Inigo-Elias et al. 2002). Sauer et al. (2003) quantified the overall species population decline to be around 55% over the last 30 years. Occurrence data also showed that the species no longer inhabits areas in south-central and eastern USA and north-eastern Mexico, where it was once abundant (Sauer et al. 1997, 2003, USFWS 2004). Historically, some estimates suggested that about 15,000 Painted Buntings per year were harvested and exported every year from Mexico up until the early 1980s (Inigo-Elias et al. 2002). In 1982, Mexico banned any wildlife exportations. Late in the 1990s, Mexico resumed the international trade of wild animals, mostly to South America, northern Europe and Malaysia (Ramos 1982; Inigo-Elias et al. 2002). Currently, and at least over the last 14 years, Mexico implemented a plan that aimed to promote and monitor sustainable harvesting of Painted Buntings and other wildlife through the implementation of harvesting quotas and geographical management units (UMAs). Unfortunately, precise control of these quotas quickly became impractical to enforce and, since the Painted Bunting and other closely related species are not listed in any of the CITES appendices, it is difficult to quantify legal and illegal captures and international exportations of any species (Cruz-Romo and Oliveras de Ita. 2011). The process by which species exploited heavily are listed in a CITES appendix and thereby gain international protection can be complex and lengthy and often require biological and ecological information that are difficult to obtain. One aspect that complicates the pathway toward an international protected status is the migratory behavior of many species of marine and terrestrial animals. Laws for the conservation

of migratory species have to take into account spatial and temporal challenges of protecting moving populations across natural landscapes and geopolitical borders. For example, the North Atlantic Right Whale (*Eubalena glacialis*) migrates great distances to specific feeding and breeding areas during its annual cycle (Cole et al. 2013). Because of these migrations, the species is subjected to different threats and level of protections at different latitudes corresponding to different national laws regarding coastal waters. On land, another well-known example of species in need of better management during migration are the wildebeest (*Connochaetes sp*.) in East Africa. Wildebeests migrate long distances out of their protected areas through migratory corridors that are often blocked by human construction (e.g. urbanization) or converted to agricultural fields (Estes and East, 2009). However, birds offer the most notable examples of long-distance migrations that entail crossing large masses of land and oceans within the same migratory journey. These migrations provide the greatest challenges in wildlife management and conservation biology. A recent study on the Northern Wheatear (*Oenanthe oenanthe*), a 25g songbird with an extensive breeding range spanning the Canadian Arctic, northern Europe, Asia, and Alaska, showed that individuals from distinct breeding populations cross the Atlantic Ocean and the Asian continent, respectively, to reach their common wintering grounds in central Africa (Bairlein et al. 2012).

Conservation biologists aiming to protect migratory species that are under human exploitation, should incorporate information from both natural populations and the market for pet trade. Unfortunately, gaining accurate knowledge of the market economic value, buyer demand, and commercial availability of the species being traded

is generally a difficult task, if international bylaws are not adopted. To enforce CITES regulations, governments and researchers implemented molecular forensic analyses to detect species illegally hunted and sold on the global market under a different species names. For example, genetic analyses showed that fish products sold in Japan included internationally protected species of whales imported or hunted illegally (Baker and Palumbi 1994). DNA analyses often proved to be effective in detecting the geographical area of origin of the individuals being illegally sold (Schwartz et al. 2007; Baker 2008; Ogden et al. 2009). The genetic structure of any animal breeding population reflects in its essence, the movements of individuals among populations, and it is a key parameter that helps to increase our understanding of species ecology and evolution (Lande 1998; Frankham et al. 2002). Modern molecular markers such as microsatellites, mitochondrial DNA, and Single Nucleotide Polymorphisms (SNPs) for example, are often useful to reveal gene flow among populations of migratory species (Avise 2004, 2008; Rundel et al. 2013; Ruegg et al. 2014). That is, genetic markers can be used to detect the influx of individuals from distant populations into the breeding population under investigation.

In this paper, I present results of a broad scale sampling, followed up by a genetic forensic investigation conducted on the Painted Bunting, as a case study of a migratory species of songbird currently sold in various pet markets in Mexico (Hamilton 2001; Inigo-Elias et al. 2002; Mlodinow and Hamilton 2005) and in Europe, and particularly valued for the adult male's brightly-colored plumage. In 2004, this species was proposed for inclusion in CITES –Appendix II, but it did not meet listing criteria (under Resolution Conf. 9.24) because the annual harvested and exported stock

from Mexico seemed, at that time, to represent no more than 0.3% of the estimated species global population (IUCN Species Survival Commission 2004; TRAFFIC, 2004). However, although Painted Buntings are abundant throughout most of their range, some small and isolated breeding populations, particularly along the U.S. east-coast, might be vulnerable to overexploitation. Here, I present a novel genetic framework that can be used to investigate the population of origin for Painted Buntings, and closely related species of the genus *Passerina*, sold worldwide. I hope to contribute toward a better understanding of which of the breeding populations of Painted Buntings are being harvested at their wintering grounds in Mexico and Central America. My genetic approach and forensic analysis will add a valuable tool for the management and conservation of multiple species of the genus *Passerina* with a particular emphasis on *P. ciris*.

Methods

Painted Bunting Sample Dataset

I analyzed genomic DNA from 138 Painted Buntings sampled at 15 sites across the species' breeding and wintering geographic range. To this panel, we added five Painted Bunting samples acquired from the pet market in central Mexico and Europe to perform population assignment tests (Fig. 3a, 3b, and Table 1). DNA aliquots for 30 birds previously examined by Herr et al. (2011) were available, and I purified genomic DNA for the rest of the samples using blood, muscles and feathers from museum collections and fieldwork (suppl. material). I included 16 samples from two populations in Jonson Bayou and Bayou Cocodrie, Louisiana, although our field site at Bayou Cocodrie was practically contiguous to the Mississippi state border (Fig 1); therefore, I refer to the

samples collected in Bayou Cocodrie as “Mississippi border population” for a better geographical characterization and to distinguish them from samples in southwestern Louisiana. Eight birds were captured at sites < 50km apart at two field stations in Costa Rica (Tamarindo and Guardia), so I treated them as if they were obtained from the same site. The dataset that I used in the program STRUCTURE was modified to investigate population structures at a regional level, and I considered the samples obtained from two populations, North Carolina and Georgia, respectively, as a representative breeding population of the Atlantic Coast area. The dataset I used in the program BOTTLENECK also had some populations merged together to investigate broad-scale past demographic fluctuations. In particular, in addition to single populations (Oklahoma, Sinaloa, and Yucatan), I included and analyzed three population clusters. These clusters were composed of samples from: Nicaragua, Costa Rica, and Guatemala (*cluster1*); eastern Texas, Arkansas, Louisiana and Mississippi (*cluster2*); North Carolina and Georgia (*cluster3*). The six samples collected in Oaxaca and the six samples collected in western Texas were excluded from the analysis due to small sample sizes.

DNA Extraction

I stored ~40ul whole blood from the brachial vein in 0.4 ml of Queen’s lysis buffer at 4° C. Muscle tissues were freeze-dried and stored at room temperature. Primary or tail feathers (first or ninth primary feather or first and second rectrices) were stored at room temperature in paper envelopes to absorb humidity and allow the feather calamus halves to dry, limiting DNA degradation. I purified DNA from blood and muscles using QIAGEN DNA extraction kit according to the protocols supplied by the

manufacturer, whereas DNA from feathers was purified following a modified feather digestion procedure (12h at 57 °C) and DNA extraction mix with no dithiothreitol (modified from De Volo et al. 2008).

Microsatellite Genotyping

For the microsatellites library, I used 13 primer sequences developed by Contina et al. (in prep.) obtained by Next Generation Sequencing (NGS) approach (454 Life Sciences, Roche). I chose these primers based on the level of loci polymorphism detected on a preliminary small subset of samples (n = 24) that were tested for unbiased allelic richness (e.g. this test is independent from sample size, see Petit et al. 1998) in FSTAT (Goudet 2002). I implemented a microsatellite genotyping method following the M13 hybrid primer technique to reduce the costs of molecular laboratory work (Boutin-Ganache et al. 2001). For the M13 hybrid primer process, I used a combination of three primers: a hybrid forward primer with 16 bp of M13F sequence (GTAAAACGACGGCCAG) tacked on the 5' end, the 16 bp M13F dye-labeled primer for which we used the 6FAM fluorescent dye (Applied Biosystems, CA), and the reverse primer. Therefore, I was able to purchase the 6FAM fluorescent dye-labeled primer in bulk for about $150 and then use small aliquots of this dye to prepare the M13 primer mix for about $15 per primer combination. I performed PCR reactions in an Eppendorf Mastercycler under the following conditions: total reaction volume 10ql, initial denaturation step of 95 °C for 15 min followed by 25 cycles of 94 °C for 30 s; 58-59 °C for 90 s; and 72 °C for 60 s. These first sets of cycles were followed by 20 additional cycles of 94 °C for 30 s; 53 °C for 90 s; and 72 °C for 60 s. A final extension of 30 min at 60 °C was conducted to ensure the addition of a terminal adenine to all the

PCR products (Brownstein et al. 1996). Fragments were analyzed using an ABI 3130XL sequencer with an internal size standard (Genescan LIZ-600, Applied Biosystems, CA). Electropherograms were visualized and allele sizes were called using PEAKSCANNER 2.0 (Applied Biosystems, CA). Samples that failed to amplify more than one locus out of the 13 loci considered were excluded from the analysis.

Genetic Diversity and Population Structure

I used the program GENEPOP 4.2 to compute several population genetic tests implementing Markov Chain Monte Carlo (MCMC) algorithm of 10,000 iterations, 10,000 dememorization steps, and significance level $\alpha < 0.01$. First, I computed the expected and observed heterozygosity for each locus within and between populations and tested for deviation from Hardy-Weinberg equilibrium with a Fisher's exact test. Then, I tested for random distribution of individuals between pairs of populations and computed a Fisher's exact probability test of population differentiation based on allelic frequencies (Raymond and Rousset 1995, Goudet et al. 1996). I finally tested for locus specific linkage disequilibrium (Raymond and Rousset 1995, Weir 1996).

I used the program ARLEQUIN 3.5.1.3 (Excoffier et al. 2009) to compute several genetic diversity and population structure tests also with the Markov chain approach. I compared populations by computing pairwise F_{ST} distances (Weir and Cockerham 1984) with 10,000 permutation steps and significant level $\alpha < 0.01$. I also used a hierarchical island model to test for loci under selection from F_{ST} statistics (Beaumont and Nichols 1996). I used the program FSTAT (Goudet 2002) to calculate unbiased estimators of gene diversity and allelic richness per locus and sample (Nei 1987, Petit et al. 1998). Finally, I used the program ML-RELATE (Kalinowski et al.

2006) to test if the individuals collected from the pet markets were closely related as parent-offspring, full-siblings, or half-siblings.

Population Clustering and Individual Assignment Tests

I implemented a Bayesian clustering algorithm using the program STRUCTURE 2.3.2.1 (Pritchard et al. 2000) to assign genotypes to clusters (K) under the assumption of Hardy-Weinberg equilibrium (Hubisz et al. 2009). We ran two different models: in *model 1*, I assumed correlated allele frequencies and admixed populations; in *model 2*, I assumed correlated allele frequencies, admixed populations, and included the collection site locations to inform the priors of the Bayesian algorithm. In both models, the number of clusters K = 1–10 were tested by 10 replicate simulations of 100,000 Markov Chain Monte Carlo (MCMC) repetitions each, and a burn-in of 10,000 iterations. I determined the most likely value for K by plotting the log probability [L(K)] of the assignment data output over multiple ΔK runs (Evanno et al. 2005) in STRUCTURE HARVESTER (Earl 2009). Results for the best models with the highest likelihood assignment scores (Francois and Durand 2010) were merged with CLUMPP (Jakobsson and Rosenberg 2007), converted into a PostScript file with DISTRUCT (Rosenberg 2004) and visualized with Adobe Illustrator CC (Adobe 2014).

To define the likelihood of correctly assigning each individual to its population of origin, I used three methods: a Bayesian (Rannala and Mountain, 1997), a frequencies-based (Paetkau et al. 1995), and a distance-based method (Goldstein et al. 1995) as implemented in the program GENECLASS2.0 (Piry et al. 2004). I calculated the assignment likelihoods by running a MCMC simulation algorithm of 100,000

iterations and a threshold assignment probability $\alpha < 0.01$. Furthermore, I also ran a genotype assignment test based on allele frequencies for all pairs of population in ARLEQUIN following Paetkau et al. (1997) and Waser and Strobeck (1998).

Population Size

In ARLEQUIN 3.5.1, I estimated the relative population size and the divergence time between populations of unequal sizes (Gaggiotti and Excoffier 2000) and calculated the Garza-Williamson index (*M*) to infer demographic history (Garza and Williamson 2001, Excoffier et al. 2005).

I also used the program BOTTLENECK 1.2.02 (Piry et al. 1999) to test for heterozygosity excess or deficiency, which are usually indicators of past population size reduction events or population expansion after a bottleneck, respectively (Cornuet and Luikart 1997). To compute these tests, I assumed a two-phased mixed model (TPM) of microsatellite evolution over time. Microsatellites are generally assumed to evolve under a stepwise mutation model (SMM), even though many loci tend to follow, at least to some extent, an infinite allele evolution model. I implemented a coalescent process in the program BOTTLENECK that tested for expected heterozygosity distribution under the TPM model with the proportion of SMM set to 70% and variance of 30% (default parameters).

Pet Trade Investigation

To find Painted Buntings for sale in the European pet market, I performed a Google search by typing the species common names in English, Spanish, French, Italian, German, and Dutch, followed by the keywords: “trade”, “sale”, “pet store”, and

"exhibition". The species common names that I used were: Painted Bunting (English); Colorín Sietecolores (Spanish); Nonpareil (French); Papa della Luisiana (Italian); Papst-Finkenammer (German); and Mexicaanse Nonpareil (Dutch). Once I found European stores that had Painted Buntings for sale, I contacted the owners by phone or e-mail to gather more information and schedule an appointment to personally see the birds and acquire samples. I acquired one adult male Painted Bunting from a large distributor of exotic bird located near Pavia, northwestern Italy, for $650.00.

In Mexico, from October to December 2013, I acquired information on eight markets in Mexico City that were referred to me as places of bird trade. I found Painted Buntings for sell only in three of them, and tail feathers were taken from four individuals: one adult male was acquired from Mercado La Argentina, a traditional market in the western sector of Mexico City, one bird of unknown sex and age was acquired from Mercado de Xochimilco, another traditional market in southern Mexico City, and one adult male together with one bird of unknown sex and age were acquired at the Cathedral of Texcoco, a town about 40 mi northwest of Mexico City. I relied on plumage characteristics to age and sex each bird according to Pyle (1997).

Results

Molecular Diversity, Hardy-Weinberg Equilibrium, and Linkage Disequilibrium

I successfully amplified 1,794 alleles across thirteen microsatellite loci for 143 samples collected at fifteen geographically distant sites and detected high allele variation between some populations and low variation between others. Among the eight breeding populations, the number of alleles ranged from 11.1 (± 5.5) for samples collected in southwestern Oklahoma and 5.3 (± 1.9) for samples collected in eastern

Texas, and the allelic size range varied from 46.9 (± 94.1) for samples collected in southwestern Texas and 25.3 (± 27.6) in eastern Texas. Samples collected in Mississippi, Arkansas, western Texas, and Oklahoma, had the highest values of genetic diversity (GD_{MS} = 0.82; GD_{AR} = 0.80; GD_{TXw} = 0.79; GD_{OK} = 0.77) and also the highest values of allelic richness ($AL(r)_{MS}$ = 1.82; $AL(r)_{AR}$ = 1.79; $AL(r)_{TXw}$ = 1.79; $AL(r)_{OK}$ = 1.77). By contrast, samples collected in eastern Texas, Louisiana, North Carolina, and Georgia, had the smallest values of genetic diversity (GD_{TXe} = 0.74; GD_{LA} = 0.75; GD_{NC} = 0.76; GD_{GA} = 0.76) and slightly lower values of allelic richness ($AL(r)_{TXe}$ = 1.73; $AL(r)_{GA}$ = 1.75; $AL(r)_{LA}$ = 1.75; $AL(r)_{NC}$ = 1.76). Among the six wintering populations, the number of alleles ranged from 10.8 (± 3.8) for samples collected in Sinaloa and 5.6 (± 2.2) for samples collected in Guatemala, while the allelic size range varied from 47.3 (± 46.2) for Sinaloa and 22.7 (± 17.1) for Guatemala. Sinaloa and Guatemala were also the two populations that showed the highest degree of genetic diversity and allelic richness (GD_{SIN} = 0.80 and $AL(r)_{SIN}$ = 1.79; and GD_{GT} = 0.79 and $AL(r)_{GT}$ = 1.79, respectively), whereas the two populations that showed the lowest degree of genetic diversity and allelic richness were Yucatan and Costa Rica (GD_{YUC} = 0.77 and $AL(r)_{YUC}$ = 1.77; and GD_{CR} = 0.78 and $AL(r)_{CR}$ = 1.77). However, I noted that the lowest genetic diversity and allelic richness values for these last two wintering populations were not statistically different when compared to the lowest values from the breeding populations in Louisiana, North Carolina, and Georgia (P_{GD} = 0.6; and $P_{AL(r)}$ = 0.5, computed in FSTAT, respectively).

Furthermore, I did not detect departure from the Hardy-Weinberg equilibrium in any of the loci from samples collected in Arkansas, eastern or southwestern Texas,

Mississippi, North Carolina, and Guatemala. However, I detected deviation from Hardy-Weinberg equilibrium in one locus from samples collected in Louisiana, Georgia, Yucatan, Nicaragua and Costa Rica, two loci in Oaxaca, and four loci in Sinaloa and Oklahoma. I did not find any loci to be in linkage disequilibrium across the fourteen populations considered.

Population Differentiation and Structure

I obtained the first genetic population structure plot for the Painted Bunting across most of its breeding and wintering range in the US and Central America using a novel library of microsatellite DNA markers. I found significant pairwise F_{ST} distances (Table S1, Supp. material) in 48% of all comparisons computed in ARLEQUIN 3.5.1 (after I removed loci in HW disequilibrium) with larger values generally observed between populations of more distant sites. F_{ST} distances ranged from 0.008 to 0.117 across all the populations from the breeding and wintering grounds representing a high variable signal of population admixture (e.g. high number of migrants) or population differentiation, respectively (Table xx suppl. material). The highest F_{ST} value was between Sinaloa and North Carolina (F_{ST} = 0.117, P < 0.00001) whereas the lowest was between Oklahoma and Sinaloa (F_{ST} = 0.008, P = 0.10), the latter of which is known to host birds from Oklahoma in the fall (Contina et al. 2013). Interestingly, the population from Georgia differed from the population in Guatemala (F_{ST} = 0.07, P = 0.0107) whereas the population from North Carolina did not differ when compared with the same population from Guatemala (F_{ST} = 0.01, P = 0.0585).

Fisher's exact probability test of population differentiation based on allelic frequencies did not reject the null hypothesis of alleles being derived from the same

distribution in 72.5% of the population pair comparisons across thirteen loci (Table xx, suppl. material). Remarkably, the two Atlantic populations and the Yucatan population were the only three populations that showed significantly different allele frequencies in all the population pairwise comparisons. Eastern Texas showed the lowest differentiation in allele frequencies.

The Bayesian population clustering algorithm computed in STRUCTURE with two slightly different model parameters yielded almost identical results (data not shown). Here, I present the results from *model 2*, which included the sample location data to inform the priors of the Bayesian clustering. Because the clustering algorithm used in STRUCTURE is sensitive to HWE deviations, I removed loci that violated this assumption before beginning clustering computations. For *model 2*, the most likely hierarchical group was found to be equal to two according to the Evanno et al. (2005) ΔK method which takes into account the mean values and the standard deviation of the posterior assignment Ln-probabilities for multiple runs (Fig 2). I also considered results obtained for K = 3, since the mean values of the posterior assignment Ln-probabilities for this clustering parameter were actually higher than the probabilities computed for K = 2, even though the standard deviation for K = 3 was slightly larger (mean LnP(K2) = -7588.2 ±20.5; and mean LnP(K3) = -7556.8 ±39.8; Fig 2). Furthermore, the results for K = 3 are useful to visualize the substructure occurring in the eastern and wintering populations (Fig 3b).

STRUCTURE results, either for K = 2 or K = 3, seem to suggest division between the western and eastern populations (Fig 3). When considering K = 3, it is even more evident that this separation take into account a further population subdivision into

three main groups: a western cluster that includes Oklahoma, Sinaloa and southern Mexico, a central cluster that includes Texas, Arkansas, Louisiana and Mississippi, and an eastern well-differentiated group that includes the populations along the Atlantic coast (Fig 3a, 3b).

Samples collected in Nicaragua and Costa Rica and on the Yucatan peninsula in winter showed a genetic mixture of birds from the large western population and birds from the Atlantic coast, either for K = 2 and K = 3 (Fig 3a, 3b, and Suppl. material). This is evidence that eastern and western populations might come into contact on the wintering grounds, where all harvesting is concentrated. For this reason, these regions should be considered targets for focusing conservation efforts.

Pet Trade Genetic Signature and Assignment Tests

Once I determined the overall genetic population structure across the species range as defined by the F_{ST} statistics computed in ARLEQUIN and the Bayesian analysis in STRUCTURE, I added to our dataset four Painted Bunting samples acquired from the pet market in Mexico and I found statistically significant differentiation between these samples and the two Atlantic populations of Painted Buntings sampled in North Carolina (FST = 0.09, P = 0.0005), and Georgia (FST = 0.11, P = 0.001).

When samples from the pet market (both Mexico and Europe) were included into the Bayesian population clustering analysis for K = 2 in the program STRUCTURE, they were regarded almost identical to samples from Texas, Arkansas, Louisiana, and Mississippi (probability of cluster membership $Q_{PET} > 0.95$)

Similarly, when these pet trade samples were included into the analysis for K = 3, they were placed into the central cluster, indicating a possible origin from the same breeding

populations listed above or from the wintering population in Guatemala (probability of cluster membership Q_{MX} = 0.9 and Q_{EU} = 0.7, respectively).

The individual assignment tests computed in GENECLASS 2.0 across all the breeding and wintering populations showed that the Bayesian assignment algorithm outperformed the allele frequency method and the distance method. The Bayesian algorithm assigned 84.7% of the individuals back to the sampling populations, while the allele frequency and distance method correctly assigned 35.3% and 10.7% of the individuals, respectively. Overall, the assignment tests computed in GENECLASS to determine the population of origin of samples acquired from the pet markets both in Mexico and Europe suggest that none of our five samples of commercial origin have genotypes similar to birds from the two Atlantic populations. Furthermore, the genotype assignment test based on allele frequencies for all pairs of populations computed in ARLEQUIN also did not assign any of the samples from the pet market to the Atlantic populations. In particular, this test suggested a high probability that the sample from Italy belongs either to the eastern Texas breeding population or to the Guatemala wintering population ($\log\text{-likelihood}_{TX}$ = -56.2 and $\log\text{-likelihood}_{GT}$ = -51.6, respectively,) as opposed to the much lower probability values to be from Georgia or North Carolina ($\log\text{-likelihood}_{GA}$ = -71.2; and $\log\text{-likelihood}_{NC}$ = -69.2, respectively). Two samples from the Mexican pet trade were assigned to the Arkansas breeding population and, for the wintering ground, to Nicaragua ($\log\text{-likelihood}_{AR}$ = -63.8 and -73.8, and $\log\text{-likelihood}_{NI}$ = -61.1 and -72.0, respectively). The last two Mexican pet trade samples were assigned to the breeding populations in eastern Texas and Mississippi and the wintering populations in Oaxaca and Guatemala ($\log\text{-likelihood}_{NC}$ =

-52.5 and log-likelihood$_{MS}$ = -53.8, and log-likelihood$_{OAX}$ = -70.1 and log-likelihood$_{GT}$ = -58.1, respectively). To further examine the resolving power of the frequency-based assignment test, I also plotted the pairwise log-likelihoods assignment probabilities of each individual from each western breeding population against the log-likelihoods assignment probabilities of each individual from the two eastern breeding populations (Fig 4). These plots further show a clear distinction between the western and eastern populations, with the exception of Louisiana and Mississippi.

Finally, my relatedness tests computed in ML-RELATE showed no indication that the birds sold in the pet trade were either related as parent-offspring, full-siblings, or half-siblings.

Past Demographic Changes

I tested for past demographic changes by computing the Garza–Williamson index in ARLEQUIN and detected a bottleneck across all the populations. The Garza–Williamson index is based on the mean ratio between alleles and allele size range and for a given population at demographic equilibrium is set at $M \geq 0.68$, but for any smaller values, and a dataset of at least seven variable microsatellite loci, we can confidently presume a demographic decline occurred in the past (Garza and Williamson 2001). The highest M value that I found was for the Oklahoma population (M = 0.48, ± 019) and the lowest value was for the Oaxaca population in southwestern Mexico (M = 0.27, ± 0.18). The extent of the decrease of M is also positively correlated with the intensity and time interval of the population decline. Thus, my M values for the Painted Bunting populations show a severe historical population bottleneck across the breeding and wintering range.

However, the test for past demographic contractions computed in BOTTLENECK did not reveal heterozygosity excess in any of the single or large population groups considered (one-tailed Wilcoxon signed–rank test; P values ranged from 0.9 to 0.4). By contrast, I detected heterozygosity deficiency in Sinaloa (one-tailed Wilcoxon signed–rank test; P = 0.002) and the test for heterozygosity deficiency was almost significant in Oklahoma (one-tailed Wilcoxon signed–rank test; P = 0.013).

Pet Trade Undercover Investigation

During my internet and field investigation in Europe, I found one large aviary facility that seem to act as main suppliers of buntings and other exotic birds located in Italy and a private molecular laboratory specialized in avian genetic sexing located in Germany. I was surprised to learn that the Italian supplier uses the molecular laboratory located in Germany to determine the sex of birds whenever plumage identification is not possible and it regularly ships live exotic birds to 26 airports across 10 countries, including several major European metropolises such as Madrid, Paris, Athens, Frankfurt, and Istanbul (Table 1 suppl. material). Furthermore, during my undercover visit as potential customer to the Italian aviary facility in early March 2013, I was allowed to visit only part of the aviary and we counted about 60 Painted Buntings and similar numbers of many other closely related species such as Indigo Bunting (*Passerina cyanea*), Lazuli Bunting (*P. amoena*), Blue Grosbeak (*P. caerulea*), and Northern Cardinal (*Cardinalis cardinalis*). Surprisingly, most of the Painted Buntings were in juvenile plumage (dull green) and, when I asked how the clients could be sure of the sex of the birds, a document from the German laboratory was shown to me with the results of DNA sex testing. However, the managers of the aviary never confirmed

(officially or otherwise) the actual origin of the birds. Therefore, I cannot conclusively affirm whether birds I saw originated from a captive breeding program or from natural areas in Middle America. Nevertheless, to the best of my knowledge, captive breeding programs for Painted Buntings have low success and bunting nestlings reared in a controlled environment often fail develop colorful plumage conditions (Contina pers. comm.; Fudickar and Kelly pers. comm.).

In Mexico, this species is sold during winter months and, in most cases, only under request, which makes it impossible to assess the number of birds held by any single vendor. When the seller had the species in stock, he/she kept a few pairs under poor cage conditions (e.g. packed together in the same cage with other species of the same size). Traditionally, the birds are sold in pair, a male and presumably a female for about $150-200 MX pesos per pair. However, Arizmendi-Arriaga (2006) states that in Mercado Sonora, a well-known pet market in downtown Mexico City, the birds used to be sold individually and the males cost $200-300 MX pesos, whereas “females” (e.g. Painted Buntings in green plumage) only $100 MX pesos.

It is worth noting that all the Painted Buntings acquired or seen in the Mexican market seemed to be illegal captures because none of them carried the band that all birds captured under an authorized permit must exhibit. In particular, the Mexican government issues two types of songbird trapping permit for pet trade. One is granted to people with very low income (usually poor peasants) for subsistence purposes; the other is for registered land areas (UMA) with sustainable harvesting systems (e.g., habitat improvement practices, harvest quotas assigned according to species density, etc.). The Dirección General de Vida Silvestre (Secretary of the Environment, SEMARNAT)

reported that between 1979 and 2000 the species was continuously trapped but no cumulative records were provided on the number of birds over these three decades. Between 2001 and 2010, the controlled capture of the species under the UMA system was authorized during five years: 2001-2002 from three UMAs in the states of Jalisco and Campeche (with no records on the number of birds), 2240 birds in 2002-2003 from three UMAs located from the very same states, in 2004-2005 with 344 birds of two UMAs from the state of Guerrero, and 2007-2008 with 1517 birds of three UMAs mainly from Guerrero (SEMARNAT 2013). Thus, I found that before the year 2000, the species was legally trapped mainly under the "subsistence" permits. After the year 2000, only permits under the "sustainable" form were granted for an overall total of a little more than 4 thousand Painted Buntings in 10 years, mainly from UMAs located in the Yucatan Peninsula and the southern Pacific coast.

Discussion

In this paper, I presented a comprehensive genetic population structure analysis for the Painted Bunting and used our gene flow inferences obtained by means of a new set of microsatellite DNA markers to reveal the population(s) of origin of birds sold in the international avian pet market. My genetic framework provided a relatively easy and low cost way to discern birds breeding along the Atlantic coast from birds breeding in the central United States. I estimated that our M13 primer mix approach was about an order of magnitude cheaper than buying separate fluorescent dye primers for each locus and offered a fast and reliable way to genotype birds across populations for about $15 each. This low cost approach and robust genetic resolving capability become particularly useful if we consider that the eastern-US population of buntings is declining

rapidly. The power of determining the population of origin of birds sold internationally as pets, could be used to guide a sustainable harvest, or stop a harmful practice.

Population differentiation, assignment tests, and genetic structure

The Painted Buntings analyzed in this study showed a detectable level of population differentiation in about half of the pairwise F_{ST} and Fisher's exact test comparisons. However, I also implemented a thorough assignment test approach carried out with different algorithms and computer programs, because this methodology has proved to be more robust than the F_{ST} statistics (Waples and Gaggiotti 2006). The Bayesian assignment algorithm for individual birds seemed to greatly outperform the allele frequency method and the distance method by correctly assigning 87.4% of the individuals back to their sampling populations of origin and identifying a little over 12% of the birds as migrants. However, considering that most Neotropical migratory songbirds have low site fidelity at the wintering grounds (Pyle et al. 2009), this result suggests that our Bayesian model implemented in GENECLASS might have over-predicted the assignment probabilities and allocated too many individuals back to the same populations from which they were sampled. Therefore, in view of the heterogeneity of our dataset made of eight and six breeding and wintering populations, respectively, I regarded the allele frequency-based assignment method to be the most reliable approach in relation to our analysis and samples. Even with the exclusion of the distance-based assigning method, which seemed to perform poorly, despite a long Monte-Carlo chain resampling approach, I gathered multiple lines of evidence from F_{ST} statistics, allele frequencies, and Bayesian population cluster analysis implemented in the program STRUCTURE that Painted Bunting samples from the international pet

market did not belong to populations breeding along the Atlantic coast, although we cannot rule out the possibility that many buntings sold as pets worldwide could belong to the populations breeding along the US Atlantic coast and wintering in Central America and the Caribbean. Here, my goal was to validate a low cost molecular tool to discern populations in relation to samples from the pet market and present the first genetic inferences about the migratory connectivity between breeding and wintering grounds of this species.

My genetic analysis confirms a particular migratory pattern for the western population recently inferred with an ultra-light tracking device. In 2013, the results of a two year study and experimentations of a custom-built miniaturized geologgers that weighted about 0.5 to 0.6 grams and fitted on 200 Painted Buntings from a breeding population in Oklahoma, revealed a loop migration of birds moving westward at the end of the summer toward Sinaloa (northwest Mexico), where they stopped to molt before continuing their migratory journey along the west coast of Mexico and then returning to the breeding sites in the spring along the east coast of Mexico (Contina et al. 2013). My genetic findings through the F_{ST} pairwise comparisons, assignment tests, and Bayesian population cluster analyses in the program STRUCTURE for samples collected in Oklahoma, Sinaloa, and Oaxaca (southern Mexico) support previous assumptions of a molt-migration strategy towards northwestern regions of Mexico in late summer, where birds can exploit higher biological productivity that results from monsoonal rains (Rohwer et al. 2005). Stopover sites for migrants should be closely monitored as potential conservation targets, where individuals from different breeding populations come into contact. In this regard, for the Painted Bunting, the Sinaloa region, the

northeast corner of the Yucatan peninsula, and to a lesser extent sites in Guatemala, Nicaragua and Costa Rica, could be considered target areas for prioritizing conservation efforts. These wintering areas that are connected by migrants throughout the annual cycle, showed a signal of genetic admixture for individuals from both western and eastern breeding populations (Fig 3a, 3b).

Admittedly, my search for the minimum and most plausible hierarchical number of clusters (K), in which the samples could be classified by the algorithm implemented in the computer program STRUCTURE, lacked a clear and well-defined value (Fig 2). The accurate and objective characterization of the parameter K has been problematic since this software was released in 2000, with the main shortcoming being the frequent underestimation of the K value (Kalinowski 2010). I tried to mitigate this limitation by implementing the Evanno et al. (2005) ΔK method in STRUCTURE HARVESTER (Earl 2009), running an increasing number of multiple replicate simulations and considering alternative values of K, if the mean posterior estimate of Ln-probabilities were relatively similar within each other and to the best Ln-probability value indicated by the Evanno et al. (2005) ΔK method. In my case, the best value of K was equal to 2, but alternative and possible values of K ranged from 2 to 6 as shown in Fig 2; because there was no positive or negative significant correlation between the value of K and higher posterior Ln-probabilities (adjusted $R^2 = 0.33$, $P = 0.18$), I also explored Bayesian clustering outputs for alternative values of K (e.g. K = 4; K = 5; and K = 6, data not shown) and found consistent results in terms of western-eastern population differentiation and pet trade samples not being clustered with the breeding populations along the Atlantic coast.

Genetic Signatures of Population Bottleneck and Expansion

My analyses of past demographic bottlenecks through two computationally different methods yielded conflicting results. These results should be interpreted in light of the mathematical differences of the tests adopted and together with considerations on the mutation rates of microsatellites. I found that the Garza-Williamson index of population size equilibrium (*M*) was extremely low across all the populations considered and indicated that a severe past demographic reduction event was still reflected in the genetic signature. On the other hand, all the tests for heterozygosity excess, usually associated with a past population bottleneck, were not statistically significant. This discrepancy could be explained by the fast mutation rates of microsatellites or the coalescent simulation model chosen for their evolution that allowed each population to quickly reach (heterozygosity) equilibrium, even after a severe and abrupt population size reduction (Rousset and Raymond 1995). In contrast, the demographic population index (*M*), which relies upon the ratio between the number of alleles and the allele size range, might require much longer to return to its equilibrium value because not every microsatellite mutation will increase the proportion of the ratio (Garza and Williamson 2001). Furthermore, the tests for heterozygosity deficiency, usually associated with a rapid population expansion or population substructure, were significant in the Sinaloa wintering population and almost significant in the southwestern Oklahoma breeding population ($\alpha = 0.01$; $P = 0.013$). This result is at least in partial agreement with the data published by Sauer et al. (1997, 2003), who showed a positive population size trend for Painted Buntings breading in southwestern Oklahoma and over the western part of the species range. Apparently conflicting results

from population bottleneck tests should take into account the sensitivity of the method used, the geographic scale, and the level of population substructure. A few demographic bottleneck studies in mammals have shown that past severe population reductions and low values of *M* might not be complemented with an occurrence of heterozygosity excess (Busch et al. 2007, McEachern et al. 2011, Shama et al. 2011), and my results for the Painted Bunting, correspond to findings from other species.

International and Intercontinental Undercover Investigations

My undercover field investigations revealed high variation in the purchase price of a single bunting between Mexico and Europe. In a street pet market in Mexico, adult Painted Bunting males are sold for $10 to $40 (US dollars) a pair. In Europe, a single Painted Bunting costs approximately a minimum of $75 (US dollars) to a maximum of $950 (US dollars), if the plumage characteristics are in excellent conditions. During my internet search and European field survey in northern Italy, I did not find substantial differences in economic values for birds of distinct age class and sex. On the other hand, ambulatory street markets are the common places for caged bird trade in Mexico, because pet stores rarely sell wild birds, perhaps because they are under (official) surveillance. The bird seller ("pajarero") is an ancient job with roots in pre-Hispanic traditions, which has been diminishing during the last few decades, probably because of new wildlife regulation, law enforcement, and changes in selling practices (pet stores as opposed to street vendors) and types of pets sought by modern city dwellers. Some "pajareros" have regular street market places that they visit, but most of them are completely itinerant. Nevertheless, regulated and unregulated caged bird trades are still

a common practice in the most populated city neighborhoods and in the country side of Mexico.

Conclusions

In conclusion, conservation biologists are confronted with great challenges when attempting to protect animals moving across a variety of ecosystems and geopolitical boundaries. Even though I focused on a small and colorful species of songbird that I consider a short to medium distance migrant protected under U.S. and Mexico laws (Migratory Bird Treaty Act; NOM-059-SEMARNAT-2010, respectively) it is steeply declining in some parts of its distribution range, but not at risk of imminent extinction, I encountered great challenges in gathering information on the species ecology and commercial status. However, despite the acquisition of individuals from the pet market for genetic analyses, and the subsequent unintended support to the trade, should be avoided whenever possible, I point out the necessity of analyzing additional samples focusing on DNA extracted from feathers, perhaps from other countries in Europe, Southern America, and Asia to determine and document which populations of Painted Buntings are exploited elsewhere.

Furthermore, I strongly believe that it is imperative to adopt conservation laws that specifically address the population structure and the migratory behavior of species undertaking extraordinary migratory journeys upon which their survival depends. In particular, I think that more specific conservation actions at the population level would benefit the eastern Painted Bunting. I note that recognizing different subspecies or geographically isolated populations should be a priority in conservation management of

this and other migratory species, since distinct breeding populations may not receive the same influx of individuals from other areas contributing to population size and genetic diversity. In this regard, given the high genetic difference between the western and eastern populations of Painted Buntings, I would recommend the eastern population to be considered an evolutionarily significant unit (ESU, sensu Moritz 1994) in accordance with Herr et al. (2011), because this level of taxonomic designations would confer a more adequate plan for species protection and management.

Acknowledgments

I thank Ann Harris and the Molecular Biology Laboratory at the University of Oklahoma. Larry Weider, Michael Patten, Bruce Hoagland for their valuable comments on earlier drafts of the manuscript. Special thanks to Tom Smith and the UCLA group for their help with the genetic analysis. The managers of the MAPS banding station managers across Central America. I thank Tyler Michels who collected blood samples in LA and NC, and many lab and field assistants who helped over the years. Philip Vasseur for providing feather samples from Louisiana. Funding for this study was also provided by the Sutton Scholarship in Ornithology at the University of Oklahoma.

Figures and tables

Figure 1. Painted Bunting breeding and wintering distribution. Triangles indicate sampling sites

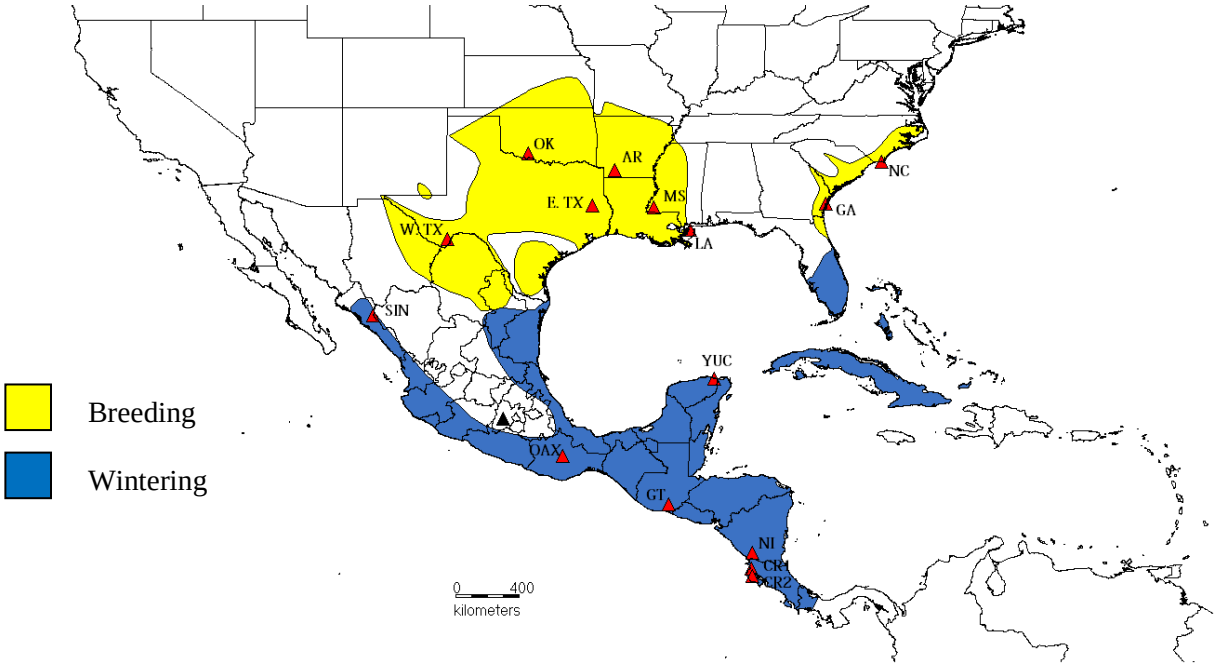


Figure 2. Posterior probabilities values indicating the most plausible number of hierarchical groups (K) in which the 14 populations can be clustered as calculated in STRUCTURE HARVESTER according to the Evanno et al. (2005) method. The posterior probabilities values were computed from 10 replicate runs for a given K.

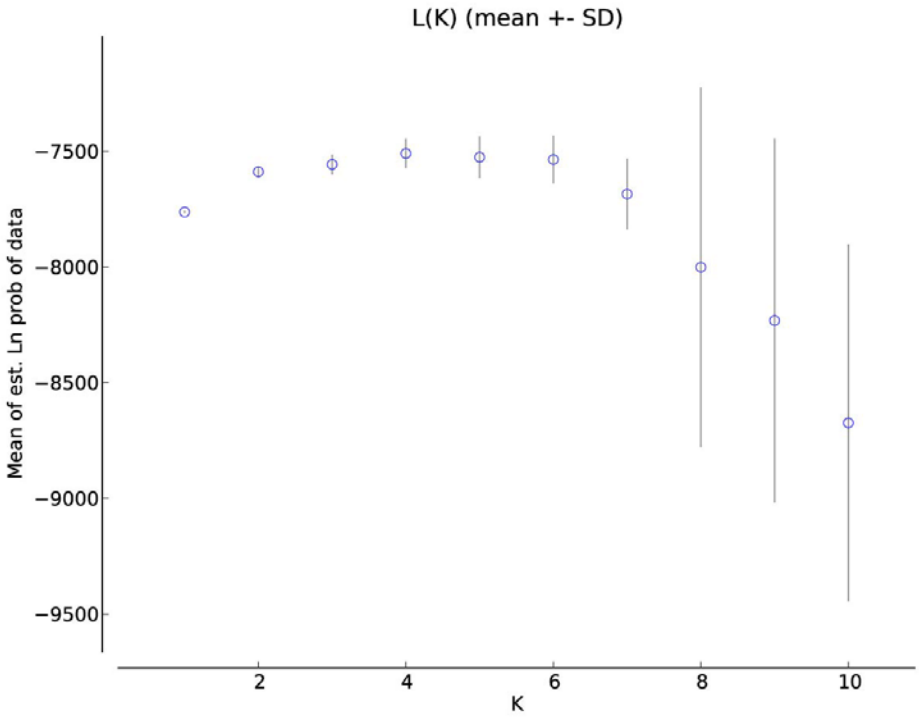


Figure 3. Population Q matrix calculated in Structure 2.3.3. The graphs show the assignment probabilities (Q) for 143 individuals from 14 populations plus the five samples from the pet trade. Each area within two vertical black lines represents a population and the area width is proportional to the number of individuals sampled in that population. The probability (Q) of each population to be assigned to a single or more clusters is shown on the vertical axis and the clusters are represented by two (for K = 2; panel 3a) or three (for K = 3; panel 3b) different colors. If a population has Q values > 0.6 is considered composed of a single and distinct cluster. The arrow and the diamond sign indicate the pet trade samples from Italy and Southern Mexico, respectively.

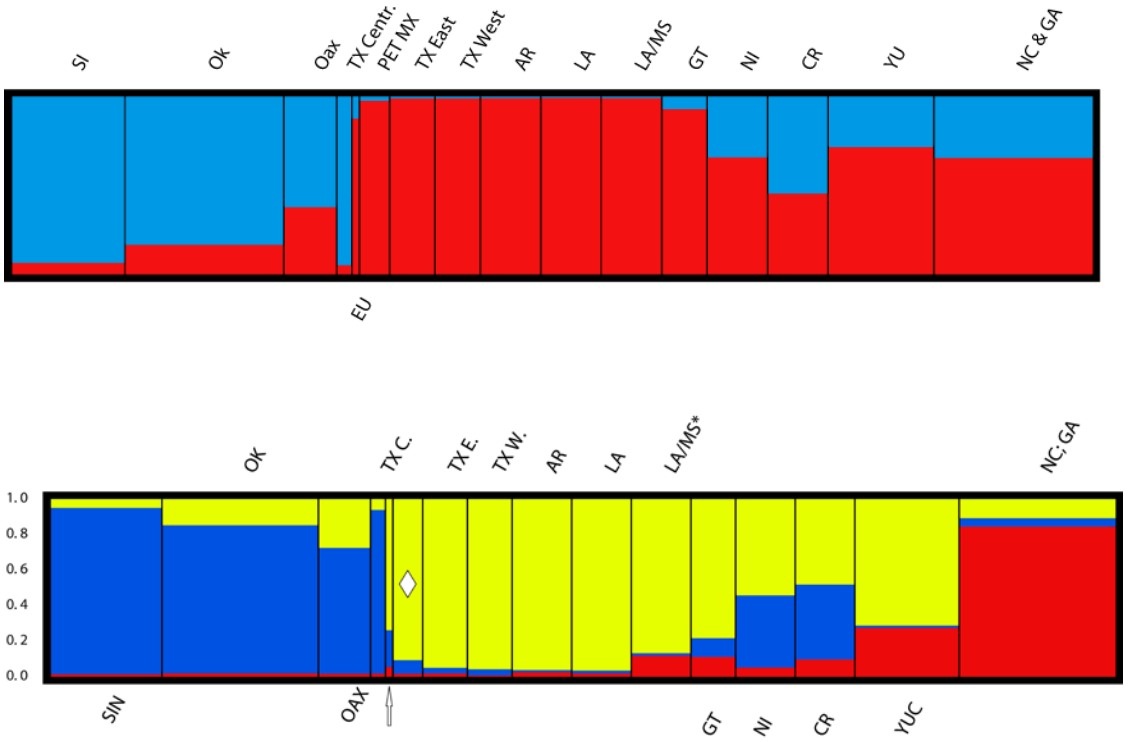
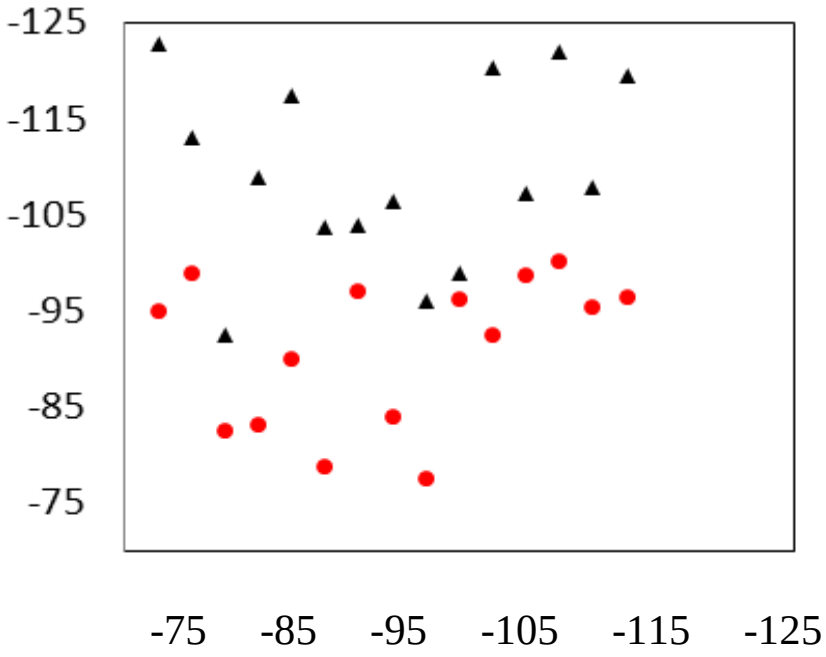


Figure 4. Log-Likelihood values of each individual being assigned to the OK population (red circles) against the Log-Likelihood values of being assigned to the NC population (black triangles).



Chapter 4: Testing for Natural Selection in candidate genes for migration in a neotropical Passerine

---

This chapter is being submitted, with some modifications, to the journal Evolution

Abstract

Complex traits such as migration are regulated by clusters of genes. Understanding how genetic variation is related to variation in migratory behavior requires both experiments with captive individuals as well as more challenging direct observations in nature. I used new ultra-light tracking devices to explore phenotypic variation in migration of a small (< 20 grams) Neotropical migrant, the Painted Bunting (*Passerina ciris*), at the intra-population level. In parallel, we initiated, through a candidate-gene approach, a search for genes that contribute to regulation of avian migration. We analyzed DNA samples from Painted Buntings (*Passerina ciris*) from three populations, one in Southwestern Oklahoma, one in Eastern Louisiana, and one in North Carolina. Birds from the Oklahoma population were tracked with geolocators allowing us to evaluate within-individual variation in behavior and genetics as well as variation across populations. I investigated possible relationships between variation in two candidate genes implicated in the control of migratory behavior (CLOCK and ADCYAP1) and several aspects of fall migration including initiation date, arrival at molt/stopover sites, and duration of the first leg of fall migration. I also evaluated the mutation rate of these candidate genes through an Approximate Bayesian Computation coalescent approach and studied the effect of natural selection on their allele frequencies across populations. I found that millennial scale summer temperatures and longitude likely affected the current genotypes of CLOCK and ADCYAP1 in my three sampling populations.

However, I found no correlation between the allele size of the microsatellite sequences within the exons of these genes and the duration or timing of fall migration. These last results disagree with a previous study on the ADCYAP1 gene in relation to migratory bird restlessness and suggest that these exonic microsatellite sequences, neither within CLOCK or ADCYAP1, can explain the variation in migratory behavior across species. Nevertheless, considering their fast mutation rate and that natural selection operates on them, CLOCK and ADCYAP1 could be candidates for population-specific adaptations during the evolution of the Painted Bunting and potentially other species of Neotropical migrants.

Introduction

A migratory life history must entail physiological and morphological adaptations that allow for sustained movement as well as behavioral adaptations that permit an animal to maintain migration in the face of uncertain weather conditions, food resources, and mortality risk (Alerstam 1990; Alerstam et al. 2003). Yet, the genetic mechanisms that underlie these physical and behavioral traits are largely unidentified. Epistatic effects, variation in gene expression, and allele size (e.g. mutation) can drastically affect animal behavior, and investigating correlations between genes and observed phenotypes can provide insight into the evolution of migration and its molecular control (Osborne et al. 1997; Ben-Shahar et al. 2002, Kitano et al. 2012). However, vertebrate genomes are vast, and it is currently not possible to compare across whole genomes (but see Gautier and Vitalis 2013). A more tenable strategy is the candidate gene approach (Kwon and Goate 2000; Wilkening et al. 2009) wherein one

focuses on a gene or a limited set of genes that are likely to relate to a particular phenotype. By studying these genes in conjunction with observations of the phenotype/s we may obtain new insight into underlying genetic mechanisms as well as evolution and adaptation of the traits of interest.

The candidate gene approach has proven to be useful in studies of genetic and phenotypic variation in several species including insects, fish, and birds (O'Malley et al. 2007, Johnsen et al. 2007, Merlin et al. 2009, Poelstra 2013). Among these attempts to identify genes associated with a particular trait, a number of studies aimed to detect genes associated with migratory phenotypes within and among populations and two prominent candidate genes emerged; *CLOCK* and *ADCYAP1*. Both genes are related to regulation of circadian rhythms in animals and their allele size polymorphisms appear to predict migratory timing and propensity for migration in some taxa (Mueller et al. 2011, Peterson et al. 2013). Nevertheless, these findings are not consistent across species and fall short of establishing clear associations between polymorphisms in candidate genes and behavioral differences outside of a captive setting.

The natural exposure to day length (photoperiod) and its seasonal variation regulates endogenous responses and migratory behaviors, but the CLOCK gene is one of the controllers for those responses tied to the circadian rhythms. From an evolutionary point of view, it is interesting to note that this gene is highly conserved across vertebrate taxa even though it is a large sequence containing 24 exons and spanning about 100,000 base pairs (100 Kb). The ADCYAP1 also acts as a controller of the circadian rhythm responses through secretion of melatonin, a hormone that regulates sleep cycles. This gene encodes a bioactive neuropeptide (adenylate cyclase activating

polypeptide) that has many other important biological functions such as insulin secretion, dilatation of blood vessels and bronchial tubes, cell differentiation, and immune system modulation (Hosoya et al. 1992; Vaudry et al. 2009). ADCYAP1 has 5 exons and 4 introns and its sequence also appears conserved across the evolution of many taxa, which imply a major role in essential biological processes. Unfortunately, even though the molecular pathways of the CLOCK gene and the ADCYAP1 gene and their functions are relatively well known, the exact physiological linkage between allele variation at these loci and phenotype (e.g. behavior) is missing (Herzog et al. 2000). In previous investigations both genes showed allelic variation and that this variation correlated with migratory phenotypes, at least in some passerines, so it seems advantageous to further explore these possible correlations in different species to determine whether these loci can be used as predictors of migratory behaviors among populations.

The present study seeks to establish or refute the validity of the genotype/phenotype correlations associated with *CLOCK* and *ADCYAP1*. To accomplish this goal I assessed *CLOCK* and *ADCYAP1* in three populations of Painted Buntings (*Passerina ciris*), some of which were tracked with geolocators as part of a separate research effort (Contina et al. 2013). In focusing on this species I was able to consider population-level differences as well as individual differences in migratory strategies to examine the relevance of the candidate genes. Moreover, my assessment of migratory behaviors is based on observations in a natural setting, as opposed to lab-based observations such as those in the study of Mueller et al. (2011).

I also studied the evolution of the CLOCK and ADCYAP1 gene through Approximate Bayesian Computation (ABC) population coalescent analyses (Csilléry et al. 2010) and a comparison of mutation rates. Finally, I attempted to test for natural selection by investigating the observed variation of these genes in relation to environmental factors across space and time.

Methods

Study Species

The Painted Bunting is a small Neotropical migrant songbird with two breeding populations inhabiting two geographically disjunct areas in North America (Fig 1), and these breeding populations differ in their patterns of migration and molt (Thompson 1991a;1991b). At the end of the breeding season in late July, the vast majority of birds in the western breeding population migrates from the southern Great Plains of the United States to northwestern Mexico to molt, and then continues the migration to the wintering grounds in southern Mexico and Central America. (Lowther et al. 1999). In contrast, the eastern population, which breeds along the Atlantic coast from North Carolina to central Florida, molts at the breeding grounds in early fall and then migrates to the wintering sites in southern Florida, the Bahamas, Cuba, and the Caribbean with some vagrants reaching Jamaica and Haiti (Rappole and Warner 1980). Moreover, Storer (1951) documented plumage coloration and wing morphology variations between the western and eastern populations. In particular, the adult males breeding along the Atlantic coast exhibit darker red feathers in their chest and also have shorter wings compared to males of the western population. Interestingly, these coloration differences are reported when birds are sampled from eastern Texas, approximately between 96°

and 97°W longitude, and eastward through part of the western population range (Am. Ornithol. Union 1957, Paynter 1970). Also, wing measurements show increased length in individuals sampled westwards from eastern Texas, particularly between 93° and 95°W longitude (Thompson 1992).

In addition to these population-level differences in migration behavior, molt, and morphologies, I have established that there is considerable variation in migratory behavior among individuals from the same populations. Contina et al. (2013) tracked Painted Buntings from southwestern Oklahoma with miniaturized light-level geolocation data loggers (hereafter "geologgers"), and documented variation in timing and duration of fall migration. Thus, the Painted Bunting makes an ideal study species for its striking variation of the migratory and molting behavior at the intra and inter-population level.

Sample Collection

I collected a total of 60 samples from adult male Painted Buntings at three geographically distant breeding sites. The first sampling site was at the Wichita Mountains National Wildlife Refuge, southwestern Oklahoma (N34.4 – W98.4), where I collected blood samples from 19 birds also equipped with geologgers. In the present study, I analyzed DNA from these same tracked individuals in conjunction with the published migratory data. For a detailed description of the methods used for geolocator deployment and data analysis see Contina et al. (2013).

The second sampling site was in Johnson Bayou, Louisiana (N30.0 – W89.1), which is near the eastern extent of the larger western population (Fig 1) representing, at least geographically and in some aspects of its morphology, an "intermediate"

population between Oklahoma and the Atlantic coast. At this site, I obtained 21 primary or secondary feathers plucked from early migrants arriving at the breeding grounds (e.g. spring migration) over the years 2011, 2012, and 2013. The earliest migrants were captured between the 8th and 19th of April of each year surveyed and, to the best of our knowledge, represent the first arrivals of Painted Buntings at the US breeding grounds ever documented with banding data over that three year period at the Johnson Bayou banding station. The rationale for including this group of early Louisiana migrants in my analysis was to test for correlations between candidate gene polymorphisms and arrival dates at the breeding grounds (spring migration) as opposed to departure dates from the breeding grounds (fall migration) inferred with geologgers in the Oklahoma population. Finally, I collected samples from 20 adult male Painted Buntings at a third breeding site in Bald Head Island, North Carolina (N33.5 –W78.0), which is a representative site of the eastern population (Fig 1). Since Painted Buntings breeding along the Atlantic coast of the US strongly differ from the western population in their patterns of migration and molt (cite Thompson or Rohwer), these samples offer an opportunity to test for candidate gene polymorphism in relation to phenotypic variation at the inter-population level.

Genotyping Analysis

I collected ~40 ul whole blood from the brachial vein from birds fitted with a geologger in Oklahoma and each individual banded in North Carolina and stored the blood in 0.4 ml of Queen's lysis buffer at 4° C. I carried out DNA isolation and purification from blood using QIAGEN DNA extraction kit and followed the protocol supplied by the manufacturer. Primary or tail feathers were collected prior to release

and following standard procedures at the banding site at Johnson Bayou – Louisiana, then stored at room temperature in paper envelopes. I isolated DNA following a modified extraction and purification protocol including an overnight tissue digestion step with no dithiothreitol (modified from De Volo et al. 2008). I used primers previously designed to amplify microsatellite repeat regions in avian species within the genes *CLOCK* and *ADCYAP1* (Steinmeyer et al. 2009). I performed 10ul (total reaction) PCR reactions following the protocols of Peterson et al. (2013). We analyzed the PCR products on an ABI 3730 sequencer with an internal size standard (Genescan LIZ-600, Applied Biosystems, CA). I then visualized and called the gene alleles using PEAKSCANNER 2.0 (Applied Biosystems, CA).

Population Differentiation and Allele Size in Relation to Phenotypes

I used the program GENEPOP to test for deviation from Hardy-Weinberg Equilibrium (HWE) and estimated the p-values through a Markov Chain Monte Carlo (MCMC) algorithm of 10,000 iterations and significant level $\alpha < 0.01$. I computed a Fisher's exact test of population differentiation for each population pair also through a MCMC approach of 10,000 iterations and significant level $\alpha < 0.01$ (Raymond and Rousset 1995, Goudet et al. 1996). I used the program ARLEQUIN (Excoffier et al. 2009) to calculate Fst statistics among populations (Weir and Cockerham 1984) and the program FSTAT (Goudet 2002) to calculate unbiased allelic richness (Nei 1987, Petit et al. 1998). In ARLEQUIN, I also calculated the genetic distances between all pairs of populations ($\delta\mu^2$) for CLOCK and ADCYAP based on the microsatellite Stepwise Mutation Model (SMM) following Goldstein et al. (1995) and used these values to build a tree diagram with the Neighbor-joining method (Saitou and Nei 1987).

Furthermore, I implemented a Bayesian population clustering algorithm in the program STRUCTURE (Pritchard et al. 2000). I analyzed each gene separately to determine the contribution of each locus to the observed population differentiation. The parameters for our Bayesian population clustering model included assumptions on correlated allele frequencies, admixed populations, and incorporated the sample locations of origin to assist the clustering algorithm, because this additional parameter does not disrupt the validity of the clustering algorithm (Hubisz et al. 2009). I searched for the optimum number of clusters (K) by running 10 replicate iterations of 100,000 MCMC for each value of K ranging from 1 to 10. I determined the most likely K value implementing the ΔK method (Evanno et al. 2005) in STRUCTURE HARVESTER (Earl 2009). I computed the Pearson correlation coefficient between allele size for each of our candidate migration genes (CLOCK and ADCYAP1) and a set of phenotypic traits associated with migratory behavior. For the Oklahoma population I considered the following migratory phenotypic traits: (i) departure dates from the breeding ground in southwest Oklahoma in late July and early August; (ii) arrival dates at the molting ground in northwestern Mexico, Sinaloa, in late summer; (iii) duration of migration from the breeding to the molting ground.

For the Louisiana population we considered the following phenotypic traits: arrival dates at Johnson Bayou, body mass, and wing chord. I computed the p-values for each Pearson coefficient with a significance level of alpha < 0.05 to test the null hypothesis that the coefficients are not significantly different from zero. I also computed simple linear regressions and the squared correlation coefficient (R^2) to explore how much of the variability of one variable was explained by the other variable

and computed the means and observed 95% confidence intervals. Finally, I calculated the Spearman correlation coefficient (rho) to explore the correlation between two morphological variables, pectoral muscle and subcutaneous fat reserves, in relation to allelic polymorphisms in the Louisiana migratory population. To assess the pectoral muscle and subcutaneous fat scores we followed the procedures proposed by Helms and Drury (1960) and (Gosler 1991), respectively. In my datasets, muscle and fat scores were ranked from 1 to 4, as whole integer scores. I computed a Student's t-test to compare CLOCK and ADCYAP1 allele size differences between the North Carolina population and the Oklahoma population. I also computed the Student's t-test to compare allele size differences between the North Carolina population and the Louisiana population. All the statistical analyses were computed in R and XLSTAT (R Development Core Team 2005; Addinsoft 2014).

Approximate Bayesian Computations

After I obtained the genetic information and determined the level of differentiation and structure among the three geographically distant breeding populations considered in this study, I treated each population as a representative genetic unit (e.g. set of individuals) in which the alleles of the CLOCK and ADCYAP1 gene evolved over time. I determined and compared the mutation rates of these two genes it with each other. My ultimate goal was to assess whether CLOCK, or ADCYAP1, or both, could be considered candidate genes for population-specific adaptations, which is why I needed to estimate their mutation rates and test for signatures of natural selection (see next method section). Thus, I implemented an Approximate Bayesian Computation (ABC) coalescent analysis (Beaumont et al. 2002)

to investigate the rate and evolution patterns of these genes through scenarios of population divergence (e.g. expansion), admixture, and geographic isolation, in the program DIYABC 2.0.4 (Cornuet et al. 2008). I compared the posterior probabilities of six different evolutionary scenarios to explain the distribution of allele frequencies that we currently observe in the Oklahoma, Louisiana, and North Carolina populations. To assess these scenarios, I estimated the posterior probabilities of ancestral effective population size (N), divergence time (t), mutation rate (μ), and rate of admixture (r), for each population. I explored two scenarios consistent with populations diverging from each other in a founder-colonization event fashion (Fig 3, scenario 1 and 2), two scenarios in which all the populations diverged from each other simultaneously from an ancestral population (Fig 3, scenario 3 and 4), and two scenarios of population isolation and divergence that allowed for genetic admixture (Fig 3, scenario 5 and 6). We implemented the ABC by simulating 10 million microsatellite datasets based on the estimates of the mean genic diversity, pairwise FST values, classification index, $(\delta\mu)^2$, and maximum-likelihood estimates of population admixture (Choisy et al. 2004). I evaluated the posterior probability of each scenario with the direct and logistic method on the 0.1 % of simulated data set closest to the observed data set (Cornuet et al. 2008). If the two methods indicated that the results are in agreement, I showed the highest posterior probabilities of the logistic method for each scenario and the cumulative posterior probabilities of scenarios from the same category (category one: scenario 1 and 2; category two: scenario 3 and 4; category three: scenario 5 and 6).

Abiotic Factors Affecting Population Structure

The genetic diversity and population structure, as well as the lack of differentiation among populations, is the direct consequence of several factors interacting with each other: historical isolation, migration, mutation, genetic drift, and natural selection. However, separating the contribution of each factor is a difficult task. I attempted to identify at least some of the environmental factors (e.g. natural selection) that could have given rise to the present CLOCK and ADCYAP1 allele frequencies detected in the three Painted Bunting populations under analysis. I examined five factors associated with the environmental features occurring within a 250 mile radius of each population geographic area and included: average temperatures and precipitation over a recent decade (1995-1985), average temperatures extrapolated from paleoclimatic models ranging from 14,000 years ago to present, geographical latitude and longitude. I retrieved the recent meteorological data from NOAA (www.ncdc.noaa.gov/data-access) and the millennial scale summer temperatures for North America from Viau et al. (2006).

I used the program GESTE (Foll and Gaggiotti 2006) to implement a hierarchical Bayesian approach to model the potential effect of five abiotic factors on the population genetic structure of CLOCK and ADCYAP1 allele frequencies for samples collected in OK, LA, and NC. For this analysis, I used three different datasets. The first dataset included CLOCK genotypes only, the second dataset included ADCYAP1 genotypes only, and the third dataset included four neutral microsatellite loci developed for the Painted Bunting by Contina et al. (in prep) and used for a forensic investigation of this species (Contina et al. in prep). These neutral loci were chosen among a larger set of microsatellite loci based on their level of polymorphism

and because they were conforming to HWE expectations and were in linkage equilibrium. Specifically, two loci had low and the other two had high allelic richness (Table X, suppl. material). Samples genotyped at these four loci were the same collected in OK and NC, and, to increase sample size, we also included samples collected in LA but from two different sites (Lafayette and Bayou Cocodrie). Using neutral loci dataset as a control allowed us to validate or dismiss my results by comparing the models built upon loci under selection (CLOCK and ADCYAP1) with models built upon loci unaffected by natural selection. I predicted that no environmental variables affecting allele frequencies would be found when using a dataset that only includes neutral markers; if environmental factors are identified, then I conclude that these factors have a broad genome-level effect likely not due to selection for particular genes.

For my computations in the program GESTE, I calculated population Fst estimates following Balding and Nichols (1995) and related these values to our list of abiotic factors using Generalized Linear Models (GLM). Since my models included five factors and their interactions, I generated a total of 32 regression models (2^5) and calculated the posterior probabilities for each model through reversible jump MCMC approach (Foll and Gaggiotti 2006). I applied these posterior probabilities to determine the model that best explained the data. The parameters of the models included 10 pilot runs of 1,000 simulations to calibrate the proposal distribution executed by the MCMC and an additional burn in of 50,000 simulations. Then, I set the simulation sample size used for final estimations to 50,000 and the number of iterations between two samples (thinning interval) to 20 to reduce the autocorrelation of the data generated from the

MCMC iterations. I used the PlotGeste.exe function to plot the posterior probabilities of the models.

Results

Allelic Variation - Oklahoma Population

For the Oklahoma breeding population tagged with geologgers at the breeding ground and released for fall migration, I successfully amplified six and seven alleles at the loci CLOCK and ADCYAP1, respectively. At the CLOCK locus, allele sizes ranged from 277 to 285 base pairs (bp). The most common allele size was 277 bp occurring in 58% of the birds. At the ADCYAP1 locus, allele sizes ranged from 165 to 177 bp. The most common allele size was 169 bp occurring in 39% of the birds. I found no correlation between allele variation at the loci CLOCK and ADCYAP1 and duration of migration, departure dates from the breeding grounds in late summer, and arrival dates at the molting site in Sinaloa. At CLOCKa , the highest correlation value was in relation to departure dates from Oklahoma and the lowest correlation value was in relation to duration of migration; both values were not statistically significantly different ($r = 0.244$; $p = 0.314$ and $r = 0.019$; $p = 0.940$, respectively). At CLOCKb, the highest correlation value was in relation to arrival dates in Sinaloa and the lowest correlation value was in relation to duration of migration, both values were not statistically significant ($r = -0.159$; $p = 0.515$ and $r = -0.117$; $p = 0.635$, respectively). Overall, the average of the two loci (CLOCKa and CLOCKb noted as CLOCKav in Table XX) did not correlate with any of the phenotypic variables investigated and it was not significant (max and min value: $r = 0.083$; $p = 0.736$ and $r = -0.031$; $p = 0.899$, respectively). At ADCYAP1a, the highest correlation value was in relation to duration of migration and

the lowest correlation value was in relation to departure dates from the breeding grounds, both values were not significant ($r = 0.115$; $p = 0.640$ and $r = -0.040$; $p = 0.870$). At ADCYAP1b, the highest correlation value was in relation to departure dates from the breeding ground and the lowest correlation value was in relation to arrival dates in Sinaloa, both values were not significant ($r = -0.358$; $p = 0.132$ and $r = -0.094$; $p = 0.702$). Overall, the average of the two loci (ADCYAP1a and ADCYAP1b noted as ADCYAP1av in Table XX) did not correlate with any of the phenotypic variables investigated (max and min value: $r = -0.260$; $p = 0.283$ and $r = -0.031$; $p = 0.899$, respectively). Interestingly, the average allele size correlation value for CLOCKav and the average allele size correlation value for ADCYAP1av in relation to arrival dates at the molting ground in Sinaloa were identical (Table XX).

Allelic Variation - Louisiana Population

For the Painted Buntings that were captured at Johnson Bayou, Louisiana during the earliest days of spring migration, I successfully amplified five alleles at both loci CLOCK and ADCYAP1. At the CLOCK gene, the allele length ranged from 277 to 286 base pairs. The most common allele size was 277 bp occurring in 60% of the birds. At the ADCYAP1 gene, the allele length ranged from 165 to 173 base pairs. The most common allele size was 169 bp occurring in 40% of the birds. I found no correlation between allele size variation at the CLOCK and ADCYAP1 loci and arrival dates at the breeding grounds during spring migration. At CLOCKa, CLOCKb, and CLOCKav (average of the two loci), the correlation values in relation to arrival dates at Johnson Bayou were not statistically significant ($r = -0.177$; $p = 0.442$, $r = 0.020$; $p = 0.930$, and $r = -0.054$; $p = 0.816$, respectively). At CLOCKav, the Spearman correlation

coefficients in relation to body fat scores and pectoral muscles scores were also not significant (rho = 0.287; p = 0.205, rho = 0.138; p = 0.548, respectively). At ADCYAP1a, ADCYAP1b, and ADCYAP1av (average of the two loci), the correlation values in relation to arrival dates at Johnson Bayou were not statistically significant (r = 0.161; p = 0.485, r = -0.118; p = 0.611, and r = 0.002; p = 0.994, respectively). At ADCYAP1av, the Spearman correlation coefficients in relation to body fat and pectoral muscles were also not significant (rho = -0.039; p = 0.868, rho = 0.113; p = 0.623, respectively).

Allelic Variation - North Carolina Population

For samples collected from the North Carolina breeding population I successfully amplified two alleles and three alleles at the CLOCK and ADCYAP1 loci, respectively. At the CLOCK locus the allele length ranged from 277 to 280 base pairs. The most common allele size was 280 bp occurring in 75% of the birds. At the ADCYAP1 locus, the allele length ranged from 167 to 171 base pairs. The most common allele size was 169 bp occurring in 90% of the birds.

None of the Student's t tests showed statistically significant allele size differences between populations. I did not find statistically significant allele size differences between the OK population and the LA population at CLOCKa (t = 0.31, p = 0.7) and CLOCKb (t = 0.03, p = 0.97) or between the OK population and the NC population at CLOCKa (t = -0.60, p = 0.54) and CLOCKb (t = 1.43, p = 0.16). Similarly, I did not find statistically significant allele size differences between the LA population and the NC population at CLOCKa (t = -0.91, p = 0.36) and CLOCKb (t = 1.22, p = 0.22). I did not find statistically significant allele size differences between the

OK population and the LA population at ADCYAP1a (t = -1.83, p = 0.07) and ADCYAP1b (t = 0.75, p = 0.45) or between the OK population and the NC population at ADCYAP1a (t = -1.81, p = 0.07) and ADCYAP1b (t = 0.04, p = 0.96). Likewise, I did not find statistically significant allele size differences between the LA population and the NC population at ADCYAP1a (t = 0.33, p = 0.73) and ADCYAP1b (t = -0.92, p = 0.36).

Population Genetic Structure

None of the populations deviated from HWE at CLOCK (p > 0.01). However, the North Carolina population deviated from HWE at ADCYAP1 (p = 0.0014). The Fisher's exact test of population differentiation for each population pair showed significant differentiation at CLOCK between OK and NC (p < 0.00001) and LA and NC (p = 0.001), but not between OK and LA (p = 0.027). The same test for population differentiation showed significant differentiation at ADCYAP1 between OK and NC (p = 0.008), but not between OK and LA (p = 0.382) and NC and LA (p = 0.215). The pairwise Fst tests among populations at CLOCK were significant only for the comparisons between OK and NC (Fst = 0.117; p = 0.006) but not between OK and LA (Fst = 0.002: p = 0.29) and LA and NC (Fst = 0.065; p = 0.032). None of the pairwise Fst tests among populations at ADCYAP1 were significant: OK and LA (Fst = -0.014; p = 0.76), OK and NC (Fst = 0.012; p = 0.044), LA and NC (Fst = -0.006; p = 0.49). I note that for a less conservative significant level of α (e.g. $\alpha < 0.05$) I would have obtained the same Fst population differentiation matrix for both genes. The results of the genetic distances ($\delta\mu^2$) computed at CLOCK and ADCYAP1 and visualized on two different Neighbor-joining trees are reported in Fig 5a and 5b, respectively. These

molecular distances computed independently at both genes identified sister populations discordantly. The sister population of LA was identified to be OK at CLOCK. By contrast, the sister population of LA was identified to be NC at ADCYAP1. At CLOCK, the allelic richness index was 6, 4.9, and 2, while at ADCYAP1 the allelic richness index was 7, 4.9, and 3, for the populations in OK, LA, and NC, respectively.

The Bayesian population clustering algorithm implemented in the program STRUCTURE suggested a moderate differentiation between the OK and LA populations (western group) and the NC population when the cluster assignments were based on CLOCK genotypes (Fig 2a). However, the same analysis revealed a lack of genetic structure between all the populations when the cluster assignments were based on ADCYAP1 genotypes (Fig 2b). The analyses were conducted for K = 3 for CLOCK and K = 2 for ADCYAP1; these were the most likely values of hierarchical clusters (K) according to the ΔK method proposed by Evanno et al. (2005), respectively. However, I need to point out that samples from NC were in HW disequilibrium at ADCYAP1, and since the clustering algorithm implemented in STRUCTURE is based on the HWE assumption, I must interpret this result with caution.

ABC Most Likely Evolutionary Scenarios

The ABC analysis compared the six evolutionary scenarios that I simulated to assess the mutation rate of CLOCK and ADCYAP1 and explain their current allele frequencies in three populations of Painted Buntings. At the CLOCK and ADCYAP1 loci, the posterior probabilities did not support scenarios that included a series of independent population divergence events (IPD) at different times (Fig 3, CLOCK IPD, scenarios 1 and 2; post. prob. were 0.09 [0.09, 0.10] and 0.10 [0.10, 0.11], respectively

and ADCYAP1 IPD, scenarios 1 and 2; post. prob. were 0.07 [0.06, 0.08] and 0.06 [0.05, 0.07], respectively), nor models that included genetic admixture events (GAE) over time (Fig 3, CLOCK GAE, scenarios 5 and 6; post. prob. were 0.08 [0.08, 0.09] and 0.19 [0.18, 0.20], respectively and ADCYAP1 GAE, scenarios 5 and 6; post. prob. were 0.03 [0.02, 0.04] and 0.12 [0.11, 0.13], respectively). By contrast, the most likely scenarios of population divergence for both CLOCK and ADCYAP1 were those that assumed an early and coinciding divergence (ECD) event from a common ancestral population into three separated populations (Fig 3, CLOCK ECD, scenarios 3 and 4; post. prob. were 0.26 [0.25, 0.27] and 0.24 [0.23, 0.25], respectively and ADCYAP1 ECD, scenarios 3 and 4; post. prob. were 0.35 [0.33, 0.36] and 0.35 [0.33, 0.36], respectively). Interestingly, not only the two genes followed the same evolutionary trajectories, as indicated by the simulated scenarios of early and concurrent divergence in the evolutionary history of the three populations of the Painted Bunting, but they also had similar mutation rates: the average mutation rate at CLOCK ($\mu = 2.51x10^{-4}$ [$1.00x10^{-4}$, $7.51x10^{-4}$]) was only slightly lower when compared to ADCYAP1 ($\mu = 3.42x10^{-4}$ [$1.10x10^{-4}$, $8.70x10^{-4}$]).

Correlation between Abiotic Factors and Allele Frequencies

Based on the analyses of genetic differentiation for samples collected in OK, LA, and NC, I proceeded to assess the effect of five abiotic factors on the population genetic structure through estimates of the regression coefficients calculated for the most likely model. Even though the Bayesian clustering algorithm in STRUCTURE and the Fst tests failed to recognize a robust genetic population structure among populations at

ADCYAP1 genotypes, I kept this gene in this analysis because the Fisher's exact test showed significant differentiation at ADCYAP1 between OK and NC ($p = 0.008$).

The results from the dataset that only included the CLOCK or ADCYAP1 genotypes for the three populations considered both, and showed higher posterior probabilities for the same two models (Fig 4a, 4b). Excluding the models that only had a constant (CLOCK *M1* = constant; prob. = 0.183 and ADCYAP1 *M1* = constant; prob. = 0.23), the two models that had the highest posterior probabilities were model 17 (CLOCK *M17* = constant; paleo-temperatures. prob. = 0.096 and ADCYAP1 *M17* = constant; paleo-temperatures. prob. = 0.076) and model 9 (CLOCK *M9* = constant, longitude; prob. = 0.094 and ADCYAP1 *M9* = constant, longitude; prob. = 0.082). Models that included present temperatures and precipitation values as well as latitude had negligible posterior probabilities (< 0.05). The results from the control dataset, which only included neutral loci, showed that no best models could be clearly identified (Fig 4c). All the models considered had negligible posterior probabilities (< 0.035). My analysis shows that two abiotic factors, longitude and paleo-temperatures, likely have an effect on the CLOCK and ADCYAP1 genotypes. However, the same factors do not have an effect when a dataset of neutral loci is considered. Therefore, I conclude that longitude and paleo-temperatures played a role in shaping the current distribution of allele frequencies at the CLOCK and ADCYAP1 gene across different populations.

Discussion

I demonstrated that allele size polymorphisms at two candidate genes for migration do not correlate with migratory phenology in three populations of the Painted

Bunting. This finding contrasts with the results presented by Mueller et al. (2011), but is similar to conclusions of Peterson et al. (2013). My results suggest that microsatellite repeats within the exons of CLOCK and ADCYAP1 are not directly linked to avian migration. However, my estimates on mutation rates indicate that CLOCK and more so ADCYAP1 are fast evolving genes. Because I also have showed that natural selection operates on these genes, I can exclude the possibility that fast evolution rate is a consequence of relaxed selective constraints and consider these as candidate genes for studying population-specific adaptive evolution.

Comparisons of Mutation Rates

The variability of microsatellite mutation rates detected across species is extremely high and it can fluctuate more than 1,000-fold among different taxonomic groups. For example, Sia et al. (2000) studied the microsatellite mutation rate in nuclear DNA of the yeast (*Saccharomyces cerevisiae*) and estimated the rates of single-repeat additions as 1.2×10^{-5}. In the Tiger Salamander (*Ambystoma tigrinum tigrinum*) Bulut et al. (2008) found an average mutation rate of 1.27×10^{-3} in ten microsatellite loci and showed that, within the genome of this species, variation across microsatellite loci is ten-fold. Brinkmann et al. (1998) reported an average microsatellite mutation rate in humans (*Homo sapiens*) of 1.93×10^{-3}. In birds, Brohede et al. (2002) investigated the germline mutation rate at microsatellite loci in the Barn Swallow (*Hirundo rustica*) and reported a fast evolution rate of 1.07×10^{-2}. By contrast, Nam et al. (2010) estimated the evolution rate of avian candidate genes associated with neurological traits and found that the average mutation rates in the chicken (*Gallus sp.*) and in the Zebra Finch (*Taeniopygia guttata*) were as low as 1.91×10^{-9} and 2.21×10^{-9}, respectively. Therefore,

my mutation rate estimates for the CLOCK and ADCYAP1 genes ($\mu = 2.51x10^{-4}$ and $\mu = 3.42x10^{-4}$, respectively) seem to conform to models of rapid evolution. However, mutation rate estimates by themselves, might not be enough to establish whether or not a fast rate of evolution is under directional selection or relaxed selective constrains. Since I determined that CLOCK and ADCYAP1 gene are targeted by natural selection (see below), we can exclude relaxed selection.

Testing for Natural Selection

Uninformative single variables and lack of clarity about mechanisms that contribute to the observed genetic population structure within a species is a difficult challenge. Nonetheless, my analyses showed that the allele frequencies of microsatellites within exon regions of CLOCK and ADCYAP1 are under similar forces of natural selection: average millennial scale summer temperatures and longitude. A niche modeling approach implemented by Shipley et al. (2013) based on paleoclimate variables suggested that the suitable historical breeding range of the Painted Bunting was distributed around the same latitudes of the Gulf of Mexico during the Last Glacial Maximum. However, the paleoclimate niche model projections also showed that three main Painted Bunting populations existed at three different but well defined longitudinal ranges (Fig. 6 reproduced with permission). My results highlight the correlation of population structure with longitude rather than latitude and these results are in agreement with Shipley's et al. niche-modeling output.

Challenges of Studying Microsatellite Sequences in Relation to Phenotypes

My results show the limitations of studying allele mutations that are functionally undetermined in relation to complex behavioral traits and I recommend future studies: 1) determine why allele polymorphism at these candidate genes do not always correlate with variation in migratory traits among species and populations and 2) attempt to explain candidate gene inconsistences in light of linkage disequilibrium, non-random mating, and genetic drift. Despite these challenges and uncertainties, developing a cost-effective molecular approach that can explain behavioral variation in migrants is highly desirable. Neutral genetic markers (e.g. DNA sequences not affected by natural selection) have been extensively used to successfully distinguish individuals, populations, and species, and are relatively easy to isolate across the genome of virtually any organism (Frankham et al. 2002, Avise 2008). Similarly, sequencing candidate genes linked to short tandem repeats (microsatellites) showing functional mutations that cause the trait of interest to vary could provide a powerful tool to study the evolution of migration.

It has been demonstrated that microsatellites are not always selectively neutral sequences and can perform crucial functions in relation to modulation of gene expression (Saunders et al. 1998) and their length expansion, for example, can cause severe neurological diseases in humans (Moxon and Wills 1999). However, testing whether the microsatellite sequences are functional or not, requires investigations of their mutation rates, genome position, and evidence that confirm evolution under natural selection. Therefore, I discourage associations between allele-size polymorphisms and various migratory phenotypes without first performing extensive

tests across different species and populations within the same species whenever possible.

Future studies on gene expression in non-model migratory organisms will help to identify novel candidate genes and will also provide the opportunity to clarify whether our knowledge and assumptions on RNA transcriptions are correct in both non-model migratory and model non-migratory species. Thus, we will likely need to look in depth into other candidate genes as soon as they become available (Fitzpatrick et al. 2005, Liedvogel et al. 2011). With the identification of novel candidate genes, even though the length and number of exons and the availability of detailed knowledge of phenotypic variation in non-model species are limiting factors, it will be easier to recognize variable genomic regions linked to these candidate genes and use them as informative markers that correlate with phenotypes. Nevertheless, I believe that the *condicio sine qua non* to successfully implement candidate gene approach is that we must find genes and/or exonic microsatellite sequences that are unequivocally associated with the regulation of migration.

Conclusions

Migration is an important life history characteristic of about 5,000 species of animals from a diversity of taxa worldwide (Riede 2001, 2004; Alerstam et al. 2003). Although advances in tracking technology, informatics, and radar have generated major advances in our understanding of how migrants use the landscape (Bridge et al. 2011, Robinson et al. 2010, Chilson et al. 2012), the genes that regulate the migratory behavior and the selective pressures that have an effect on animal's migratory

adaptation remain largely unknown (Webster et al. 2005, Bowlin et al. 2010). The genes CLOCK and ADCYAP1 play a role in regulating the circadian rhythms, which are part of those endogenous programs that help any living organism to cope with the periodic characteristics of the natural environment and have been found and observed in vertebrates, invertebrates, single cells, enzymes, and in crucial molecular processes such as gene expressions (Pittendrigh 1993). Migratory species depend on such endogenous programs to determine when is time to migrate, as well as direction and duration of their movements (Berthold et al. 2003). I cannot preclude that these genes could be important for the migratory behavior but my study shows that the specific microsatellite variations within CLOCK and ADCYAP1, here and previously considered, cannot be conclusively linked to gene expression to a degree that justifies the observed migratory differences at the intra-population level in the Painted Bunting.

Acknowledgments

I thank Frank Moore, Kristen Covino and the Migratory Bird Research Group of the Department of Biological Sciences at The University of Southern Mississippi in Hattiesburg, for collecting the feather samples at Johnson Bayou, Louisiana. I thank Ann Harris and the Molecular Biology Laboratory at the University of Oklahoma. I am grateful to Larry Weider, Michael Patten, Bruce Hoagland, and Tom Smith for valuable comments that greatly improved the manuscript. Funding for this study was also provided by the Sutton Scholarship in Ornithology at the University of Oklahoma.

Figures and tables

Figure 1. Painted Bunting Distribution. The breeding and wintering ground are indicated in green and orange, respectively. The red dots indicate the sampling locations at the Wichita Mountains National Wildlife Refuge in Oklahoma (OK), Johnson Bayou in Louisiana (LA) and Bald Head Eagle in North Carolina (NC).

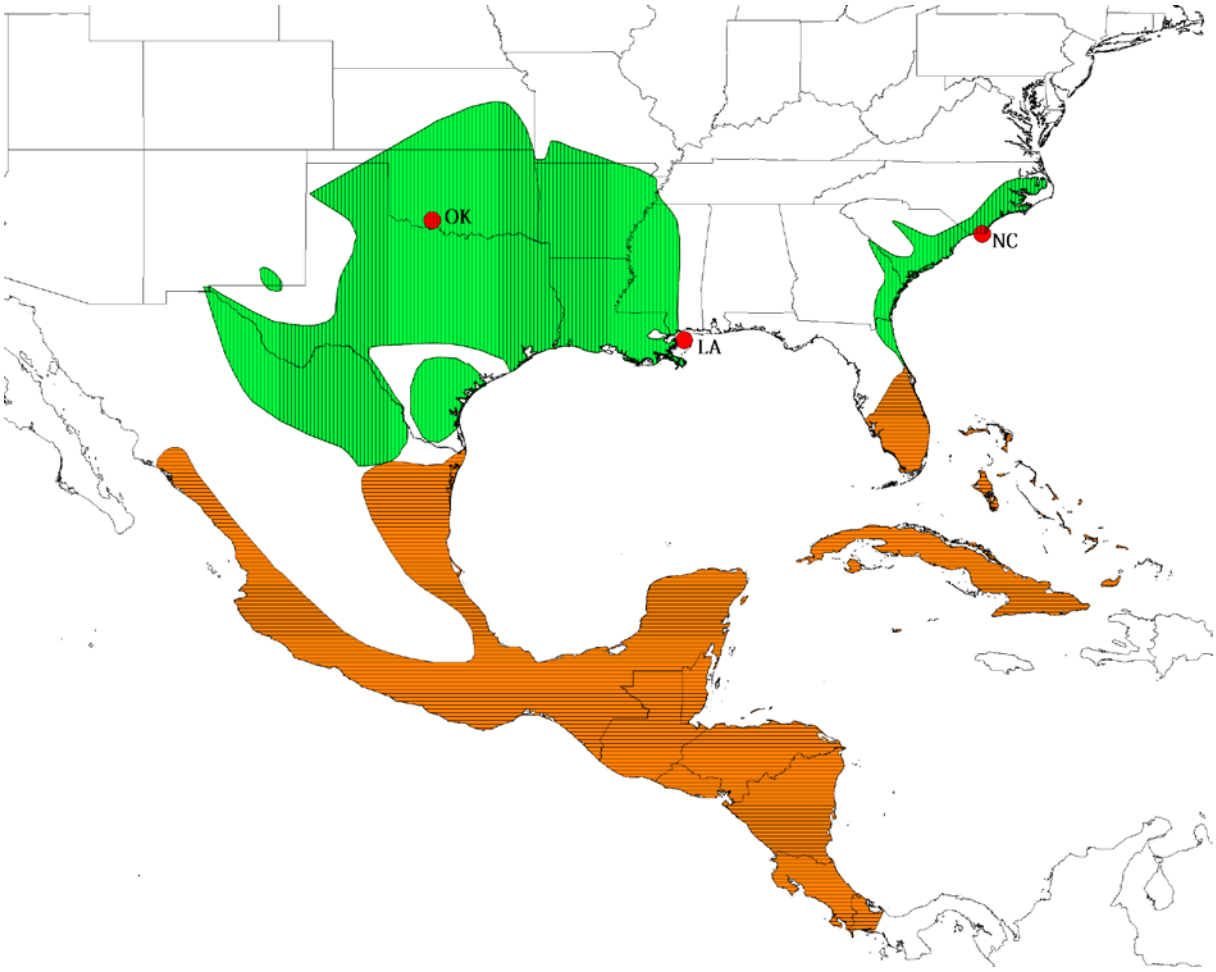


Figure 2. STRUCTURE Q matrix outputs. CLOCK gene, K = 2. The graph shows the assignment probabilities (Q) for 60 individuals from 3 populations (OK, LA, and NC). The probability (Q) of each population to be assigned to a single or more clusters is shown on the vertical axis and the clusters are represented by two different colors.

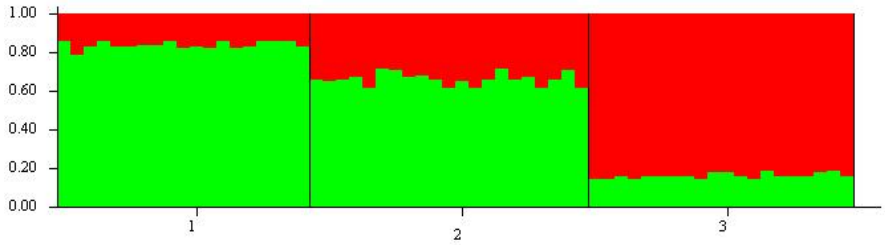


Figure 3. ABC Scenarios. Oklahoma (Pop 1), Louisiana (Pop 2), North Carolina (Pop 3). Scenario 3 and 4 have cumulative posterior probabilities of 0.5 (logistic method).

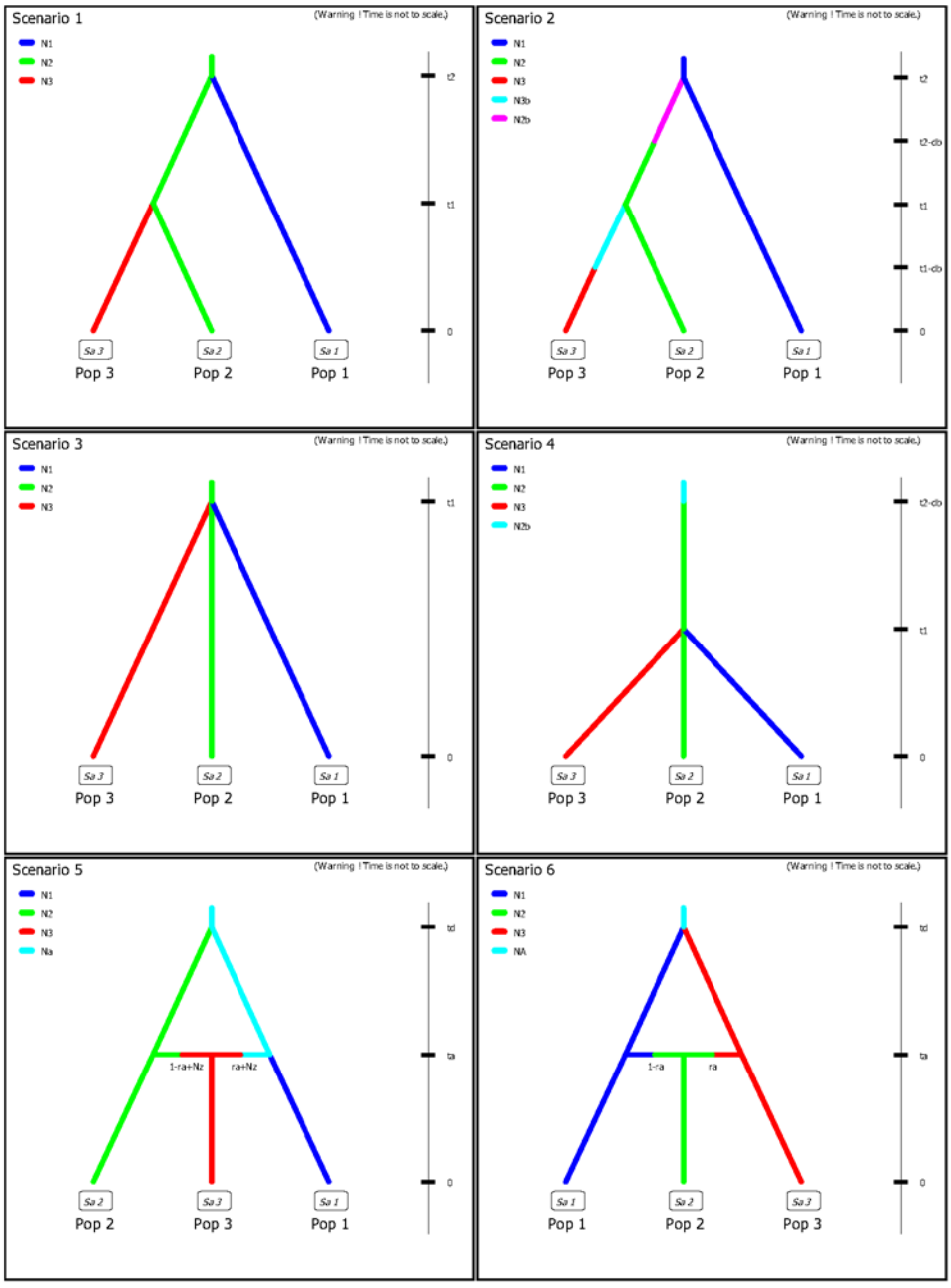
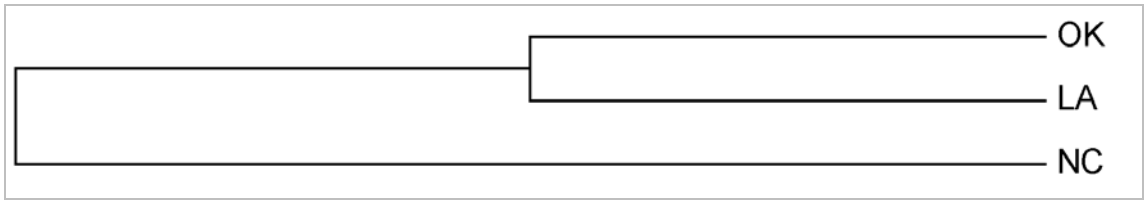


Figure 4 GESTE results. Posterior Model Probabilities – Only CLOCK, ADCYAP1, and neutral loci.

Figure 5a. The results of the genetic distances ($\delta\mu^2$) computed at CLOCK and visualized on a Neighbor-joining tree. Oklahoma and Louisiana are sister populations. Branches are not on scale. Fig 5b. The results of the genetic distances ($\delta\mu^2$) computed at ADCYAP1 and visualized on a Neighbor-joining tree. Louisiana and North Carolina are sister populations. Branches are not on scale.



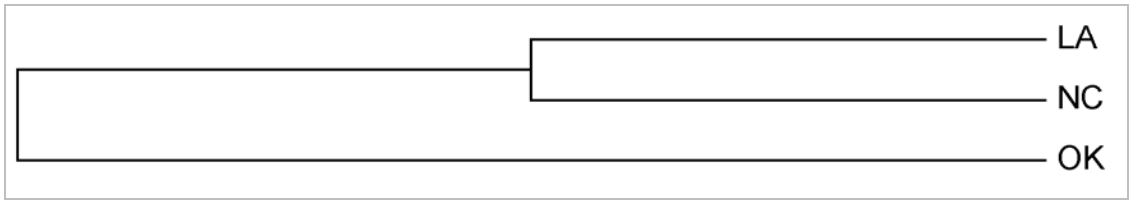


Figure 6. Painted Bunting Paleo-climate breeding models. Modified from Shipley et al. 2013.

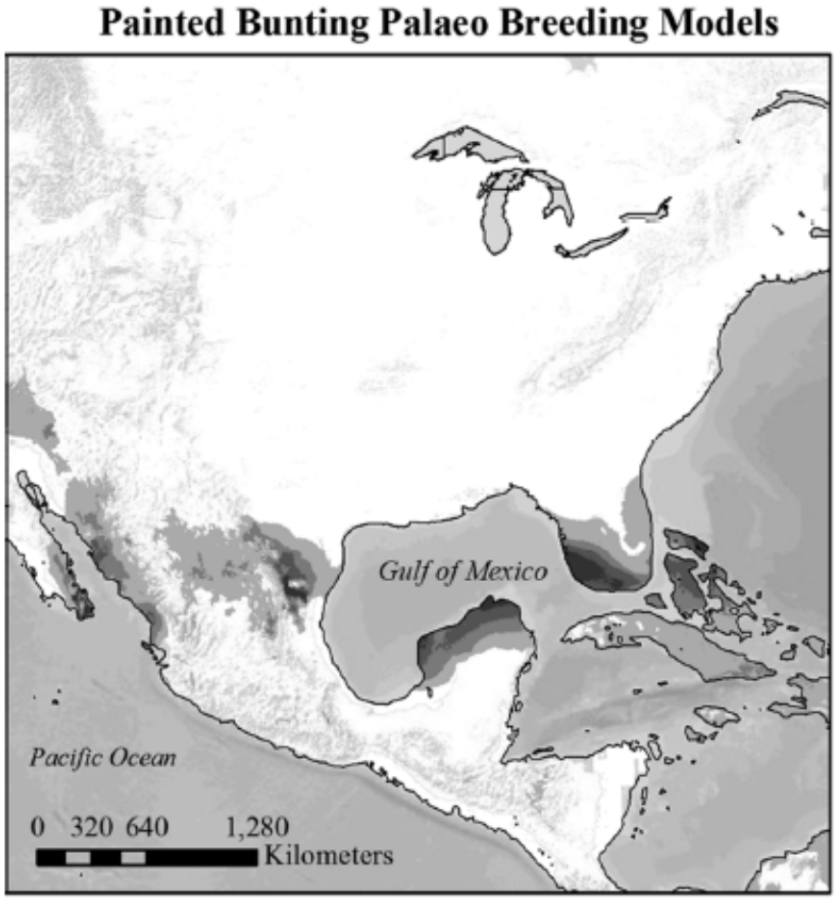


Table 1. Oklahoma CLOCK and ADCYAP1 genotypes

Band ID	Location	CLOCK-a	CLOCK-b	CLOCKav	ADCYAP1-A	ADCYAP1-B	ADCYAP1av
222150751	Oklahoma	277	285	281	167	171	169
222150752	Oklahoma	277	280	278.5	167	169	168
222150754	Oklahoma	280	283	281.5	165	169	167
222150755	Oklahoma	277	283	280	169	177	173
222150757	Oklahoma	277	277	277	167	169	168
222150758	Oklahoma	277	277	277	167	171	169
222150760	Oklahoma	280	282	281	167	170	168.5
251112762	Oklahoma	280	282	281	167	169	168
251112838	Oklahoma	277	283	280	167	171	169
251112852	Oklahoma	277	277	277	169	171	170
251112854	Oklahoma	277	277	277	167	169	168
251112858	Oklahoma	277	277	277	167	169	168
251112859	Oklahoma	277	283	280	169	171	170
251112861	Oklahoma	277	277	277	169	171	170
251112864	Oklahoma	277	277	277	169	173	171
251112865	Oklahoma	277	283	280	167	169	168
251112866	Oklahoma	274	277	275.5	169	170	169.5
251112883	Oklahoma	277	283	280	165	169	167
251112890	Oklahoma	280	283	281.5	169	170	169.5

Table 2. Louisiana CLOCK and ADCYAP1 genotypes

Band ID	Location	CLOCK-a	CLOCK-b	CLOCKav	ADCYAP1-A	ADCYAP1-B	ADCYAP1av
251151057	Louisiana	277	277	277	169	173	171
192145897	Louisiana	277	277	277	167	169	168
192145899	Louisiana	277	277	277	169	171	170
192172001	Louisiana	280	286	283	167	171	169
192173293	Louisiana	277	280	278.5	169	169	169
192172009	Louisiana	277	283	280	169	169	169
192172012	Louisiana	277	286	281.5	167	171	169
192172019	Louisiana	280	286	283	169	169	169
192173040	Louisiana	280	283	281.5	169	169	169
192172020	Louisiana	277	277	277	167	167	167
192172022	Louisiana	277	280	278.5	169	173	171
251151096	Louisiana	277	277	277	169	169	169
251151100	Louisiana	277	280	278.5	165	171	168
251151101	Louisiana	277	277	277	167	167	167
192172072	Louisiana	271	277	274	171	171	171
192172076	Louisiana	277	277	277	167	167	167
192172085	Louisiana	280	283	281.5	169	173	171
192173061	Louisiana	277	280	278.5	167	167	167
192173080	Louisiana	277	277	277	171	171	171
251151168	Louisiana	277	286	281.5	169	171	170
251151173	Louisiana	277	280	278.5	169	171	170

Figure 7. Regression models for the Oklahoma population

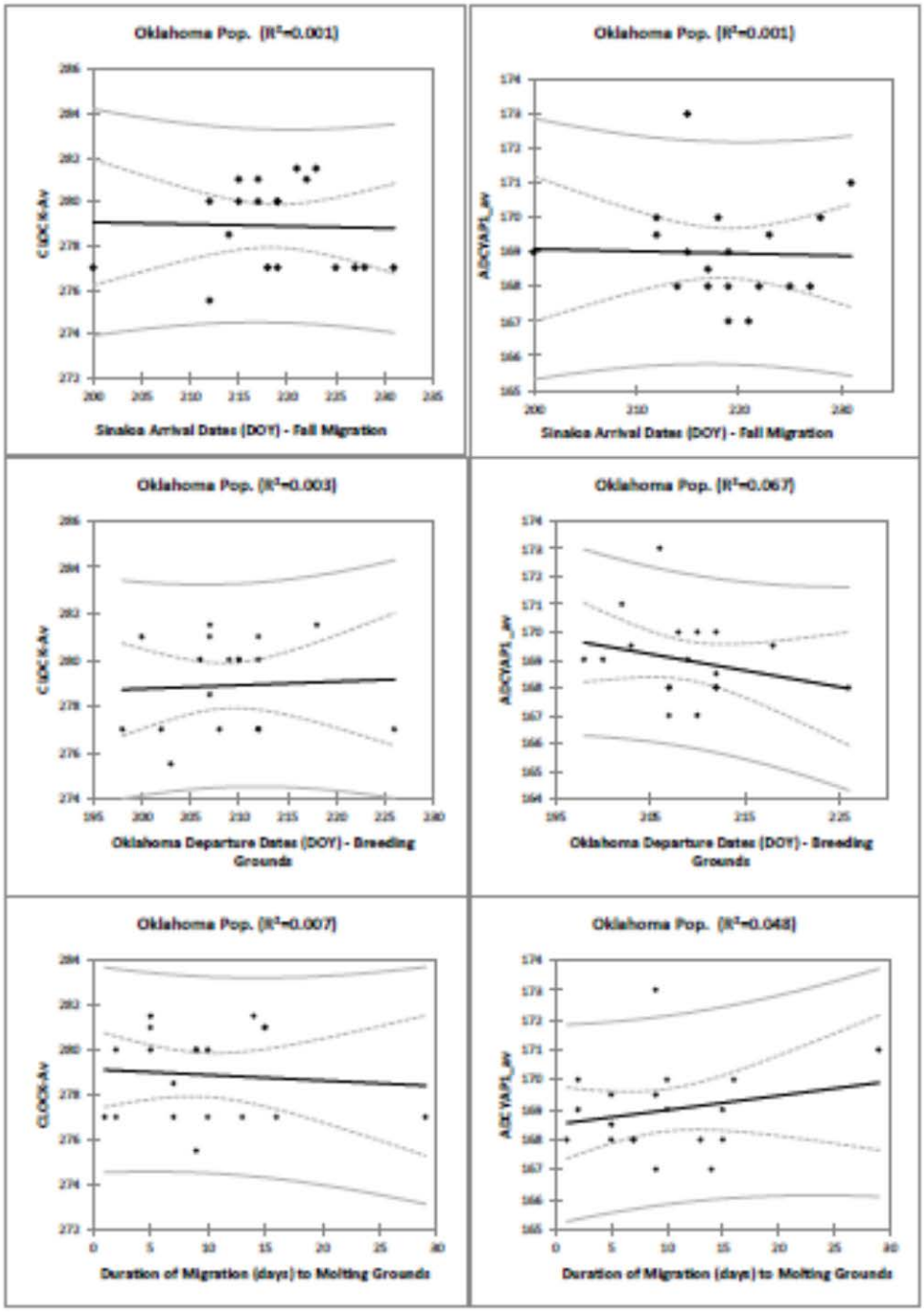


Figure 8. Regression models for the Louisiana population

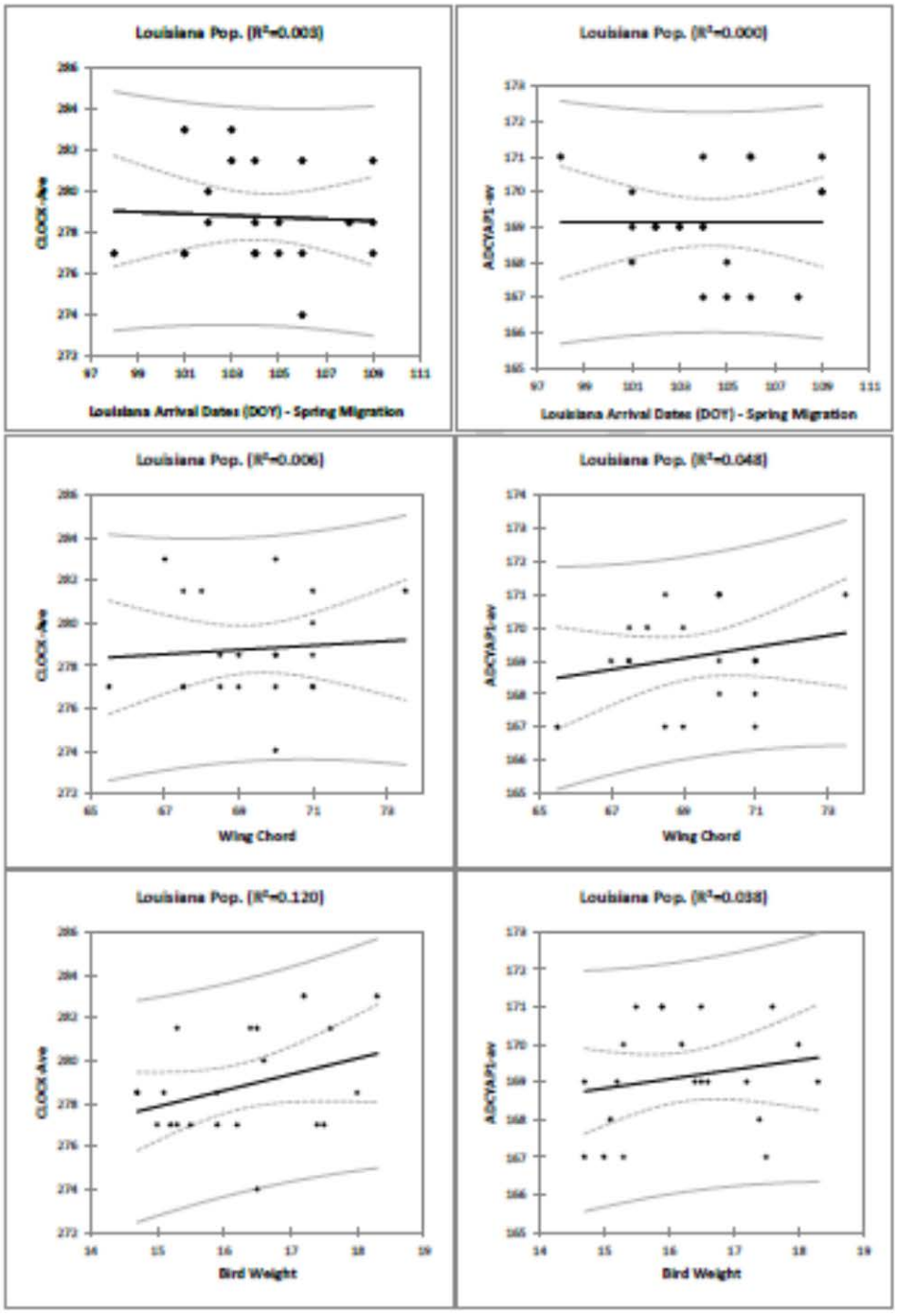


Table 3. Correlation matrix for the Oklahoma population

Variables	Start DOY	Arrival DOY	duration	CLOCK-a	CLOCK-b	CLOCK-Av	ADCYAP1-A	ADCYAP1-B	ADCYAP1_av
Start DOY	**1**	**0.519**	-0.416	0.244	-0.055	0.051	-0.040	-0.358	-0.260
Arrival DOY	**0.519**	**1**	**0.561**	0.240	-0.159	-0.031	0.071	-0.094	-0.031
duration	-0.416	**0.561**	**1**	0.019	-0.117	-0.083	0.115	0.247	0.218
CLOCK-a	0.244	0.240	0.019	**1**	0.440	**0.721**	-0.303	-0.185	-0.262
CLOCK-b	-0.055	-0.159	-0.117	0.440	**1**	**0.939**	-0.296	0.089	-0.074
CLOCKav	0.051	-0.031	-0.083	**0.721**	**0.939**	**1**	-0.344	-0.002	-0.157
ADCYAP1-A	-0.040	0.071	0.115	-0.303	-0.296	-0.344	**1**	**0.562**	**0.830**
ADCYAP1-B	-0.358	-0.094	0.247	-0.185	0.089	-0.002	**0.562**	**1**	**0.928**
ADCYAP1av	-0.260	-0.031	0.218	-0.262	-0.074	-0.157	**0.830**	**0.928**	**1**

Values in bold are different from 0 with a significance level alpha=0.05

Table 4. Oklahoma population. Correlation p-values:

Variables	Start DOY	Arrival DOY	duration	CLOCK-a	CLOCK-b	CLOCK-Av	ADCYAP1-A	ADCYAP1-B	ADCYAP1_av
Start DOY	**0**	**0.023**	0.076	0.314	0.824	0.835	0.870	0.132	0.283
Arrival DOY	**0.023**	**0**	**0.012**	0.323	0.515	0.899	0.772	0.702	0.899
duration	0.076	**0.012**	**0**	0.940	0.635	0.736	0.640	0.308	0.369
CLOCK-a	0.314	0.323	0.940	**0**	0.060	**0.000**	0.207	0.448	0.279
CLOCK-b	0.824	0.515	0.635	0.060	**0**	**< 0.0001**	0.218	0.718	0.764
CLOCK-Av	0.835	0.899	0.736	**0.000**	**< 0.0001**	**0**	0.149	0.993	0.521
ADCYAP1-A	0.870	0.772	0.640	0.207	0.218	0.149	**0**	**0.012**	**< 0.0001**
ADCYAP1-B	0.132	0.702	0.308	0.448	0.718	0.993	**0.012**	**0**	**< 0.0001**
ADCYAP1_av	0.283	0.899	0.369	0.279	0.764	0.521	**< 0.0001**	**< 0.0001**	**0**

Values in bold are different from 0 with a significance level alpha=0.05

Table 5. Correlation matrix for the Louisiana population

Variables	Date	Fat Score	Muscle	W. Chord	Mass	CLOCK-av	CLOCK-a	CLOCK-b	ADCYAP1-A	ADCYAP1-B	ADCYAP1-av
Date -	**1**	**-0.697**	0.396	0.026	-0.101	-0.054	-0.177	0.020	0.161	-0.118	0.002
Fat Score	**-0.697**	**1**	-0.251	0.031	**0.520**	0.384	0.268	0.379	-0.153	0.114	0.000
Muscle	0.396	-0.251	**1**	0.047	0.194	0.178	0.347	0.057	0.136	0.040	0.097
W. Chord	0.026	0.031	0.047	**1**	0.303	0.078	0.122	0.041	0.224	0.151	0.219
Mass	-0.101	**0.520**	0.194	0.303	**1**	0.346	0.299	0.312	0.074	0.226	0.194
CLOCK-av	-0.054	0.384	0.178	0.078	0.346	**1**	**0.781**	**0.944**	-0.157	0.162	0.031
CLOCK-a	-0.177	0.268	0.347	0.122	0.299	**0.781**	**1**	**0.530**	-0.256	0.004	-0.128
CLOCK-b	0.020	0.379	0.057	0.041	0.312	**0.944**	**0.530**	**1**	-0.077	0.218	0.111
ADCYAP1-A	0.161	-0.153	0.136	0.224	0.074	-0.157	-0.256	-0.077	**1**	0.372	**0.769**
ADCYAP1-B	-0.118	0.114	0.040	0.151	0.226	0.162	0.004	0.218	0.372	**1**	**0.880**
ADCYAP1av	0.002	0.000	0.097	0.219	0.194	0.031	-0.128	0.111	**0.769**	**0.880**	**1**

Values in bold are different from 0 with a significance level alpha=0.05

Table 6. Louisiana population. Correlation p-values:

Variables	Date - DOY	Fat Score	Muscle	W. Chord	Mass	CLOCK av	CLOCK -a	CLOCK -b	ADCYAP 1-A	ADCYAP 1-B	ADCYAP1 av
Date - DOY	**0**	**0.000**	0.076	0.910	0.662	0.816	0.442	0.930	0.485	0.611	0.994
Fat Score	**0.000**	**0**	0.273	0.895	**0.016**	0.086	0.239	0.090	0.508	0.623	1.000
Muscle	0.076	0.273	**0**	0.841	0.400	0.440	0.123	0.805	0.555	0.864	0.674
W.Chord	0.910	0.895	0.841	**0**	0.182	0.737	0.599	0.859	0.328	0.513	0.340
Mass	0.662	**0.016**	0.400	0.182	**0**	0.124	0.188	0.169	0.749	0.324	0.400
CLOCKav	0.816	0.086	0.440	0.737	0.124	**0**	**< 0.01**	**< 0.01**	0.497	0.482	0.893
CLOCK-a	0.442	0.239	0.123	0.599	0.188	**< 0.001**	**0**	**0.013**	0.263	0.987	0.579
CLOCK-b	0.930	0.090	0.805	0.859	0.169	**< 0.001**	**0.013**	**0**	0.739	0.342	0.633
ADCYAP1-A	0.485	0.508	0.555	0.328	0.749	0.497	0.263	0.739	**0**	0.097	**< 0.0001**
ADCYAP1-B	0.611	0.623	0.864	0.513	0.324	0.482	0.987	0.342	0.097	**0**	**< 0.0001**
ADCYAP1av	0.994	1.000	0.674	0.340	0.400	0.893	0.579	0.633	**< 0.0001**	**< 0.0001**	**0**

Values in bold are different from 0 with a significance level alpha=0.05

Table 7. Spearman correlation matrix

Variables	Fat Score	(1)Muscle	CLOCK-a	CLOCK-b	CLOCK-Ave	ADCYAP1-A	ADCYAP1-B	ADCYAP1-av
Fat Score	**1**	-0.269	0.325	0.281	0.287	-0.122	0.080	-0.039
(1)Muscle	-0.269	**1**	0.269	0.060	0.138	0.164	0.042	0.113
CLOCK-a	0.325	0.269	**1**	**0.591**	**0.736**	-0.098	0.046	-0.032
CLOCK-b	0.281	0.060	**0.591**	**1**	**0.976**	-0.056	0.217	0.084
CLOCK-Ave	0.287	0.138	**0.736**	**0.976**	**1**	-0.102	0.179	0.035
ADCYAP1-A	-0.122	0.164	-0.098	-0.056	-0.102	**1**	0.412	**0.812**
ADCYAP1-B	0.080	0.042	0.046	0.217	0.179	0.412	**1**	**0.859**
ADCYAP1-av	-0.039	0.113	-0.032	0.084	0.035	**0.812**	**0.859**	**1**

Values in bold are different from 0 with a significance level alpha=0.05

Table 8. Spearman correlations p values

Variables	Fat Score	(1)Muscle	CLOCK-a	CLOCK-b	CLOCK-Ave	ADCYAP1-A	ADCYAP1-B	ADCYAP1-av
Fat Score	**0**	0.238	0.150	0.215	0.205	0.598	0.729	0.868
(1)Muscle	0.238	**0**	0.236	0.795	0.548	0.473	0.857	0.623
CLOCK-a	0.150	0.236	**0**	**0.005**	**0.000**	0.672	0.841	0.890
CLOCK-b	0.215	0.795	**0.005**	**0**	**< 0.0001**	0.810	0.342	0.715
CLOCK-Ave	0.205	0.548	**0.000**	**< 0.0001**	**0**	0.658	0.433	0.880
ADCYAP1-A	0.598	0.473	0.672	0.810	0.658	**0**	0.064	**< 0.0001**
ADCYAP1-B	0.729	0.857	0.841	0.342	0.433	0.064	**0**	**< 0.0001**
ADCYAP1-av	0.868	0.623	0.890	0.715	0.880	**< 0.0001**	**< 0.0001**	**0**

Values in bold are different from 0 with a significance level alpha=0.05

Chapter 5: Allele variation at the MYH7 gene correlates with speed of migration across 11 species of birds.

This chapter is being submitted, with some modifications, to the journal Molecular Ecology.

Abstract

To search for genes associated migratory performance, I implemented 454 next generation sequencing and generated over 48,000 DNA sequences randomly distributed over the Painted Bunting genome. In parallel, I assembled an extensive library of candidate genes by annotations extracted from the Mouse Genome Informatics database. I used about 30% of the 454 reads output within a candidate gene framework and identified a novel set of 12 candidate migration genes. In particular, I focused on the following candidate genes: ADRA1d, ANKRD17, CISH, and MYH7, and I tested them across 21 avian migratory and non-migratory species revealing a surprising degree of allelic variation in some genes. My validation effort was performed through the implementation of light-level geolocator (e.g. geologger) analyses and collection of DNA samples at the intra-species level in the Painted Bunting, and through extensive gathering of migration data across avian species from the published literature. I found that the polymorphism at the gene MYH7 correlated with speed of migration (km/day) and body mass across 11 species of songbirds. Furthermore, I found two novel uncharacterized polymorphic genes strongly correlating with duration of migration in the Painted Buntings.

Introduction

The correlation between genetic variation and phenotype is the basis for much of modern evolutionary, developmental, and medical biology (Wong et al. 2005; Vaysse et al. 2011). By examining genetic polymorphisms that control certain traits, we can make inferences about phenotypic variation and perhaps even make predictions about disease risks at the individual or population level (HDCRG 1993; Zatoń-Dobrowolska et al. 2014). Basic biological research and medical applications have examined thousands of genes in mice that regulate the cardiovascular system, metabolism, muscle strength, behavior, circadian rhythms, and brain activities under controlled laboratory conditions (Alberts et al. 2002). The last three decades of genetic investigations on mice have produced an impressive amount of information describing the functions of thousands of genes that have been collected and made publicly available through the Mouse Genome Database (MGD, http://www.informatics.jax.org). This effort includes a comprehensive catalog of gene/phenotype functional associations that constantly integrates new data from the National Center for Biotechnology Information (NCBI) and provides a reliable genomic resource for the international scientific community interested in using the mouse (*Mus musculus* spp.) as a model organism for broader genetic research applications (Blake et al. 2014). However, whether the genes studied in mice occur, are expressed, and have identical functions in non-model organisms are open questions (Enard et al. 2002, Liao and Zhang 2006).

In studies of avian migration, sequences of DNA known to be involved in regulation of heart morphological modifications, cellular physiology, fiber muscular strength and growth, regulation of circadian rhythms, food digestion, exploratory

behaviors, and spatial orientation, are promising starting points in the search for candidate genes that impact migratory phenotypic variation within and across migratory species (Pulido et al. 1996; Pulido 2007). Indeed, several physiological responses under genetic control, such as striking modification of supporting flight organs, hypertrophy of the heart, supracoracoid, and pectoral muscles, can influence migration in terms of onset, length, direction, and ultimately, migration success (Piersma et al. 1999; Bauchinger and Biebach 2001; Bauchinger et al. 2005). Furthermore, genes that regulate fat metabolism play an essential role in fueling the migratory journey because lipids provide a high ratio of calories per gram of body mass (Guglielmo et al. 2002). Migratory birds are able to store conspicuous lipid reserves (fat) before migration (McFarlan et al. 2009), and measurements of these subcutaneous fat reserves can provide estimates of incipient migration in birds that, in conjunction with the intensity of migratory restless condition (*Zugenruhe*), are positively correlated with distance to be travelled before reaching the stopover site or wintering ground (Loria and Moore 1990; Yong et al. 1998; Alerstam and Hedenstrom, 1998; Buler et al. 2007; Deutschlander and Muheim 2009). In addition, when birds are physiologically ready for migration, they need to orient to the exact direction and navigate themselves to their destination (Åkesson and Anders 2007). This is usually accomplished using a combination of recognizing natural or artificial landmarks, exploiting the Earth's geomagnetic field, and using celestial clues, both during the day and at night (Wiltschko and Wiltschko 2003; Berthold et al. 2003). Therefore, genes that regulate neurological synapses in areas of the brain deputed to spatial intelligence could play a role in shaping migration phenotypes.

Avian genetic research has focused on two promising candidate migration genes, CLOCK and ADCYAP1. These genes are involved in controlling human and animal circadian rhythms, which are also known to be important in determining the onset of the migratory restless condition in birds. In a recent European study (Mueller et al. 2011), it was found that the allele size of exon microsatellite sequences within ADCYAP1 correlated with migratory departure dates or predisposition for migration in the Blackcap (*Sylvia atricapilla*). Nevertheless, these correlations were not repeatable across passerines species (Peterson et al. 2013; Contina et al. in prep). Therefore, candidate genes that are unmistakably linked to variation of avian migratory behavior are yet to be discovered.

In this paper, I report my attempt to find novel candidate genes associated with different migratory behaviors in avian species. I implemented a next generation sequencing (NGS) method and highlighted the benefits and limitations of the candidate gene approach for investigations centered on a complicated and highly variable phenotypic trait such as animal migration. My findings indicate an interesting variation in microsatellite sequences within some important functional genes, but stress the necessity of performing intra- and inter-species validation to conclusively accept or refute the correlations between high allelic polymorphism and phenotypic variation of a given trait.

Methods

The objective of this study was to identify and validate new candidate migration genes at the intra- and inter-species level. My intraspecific study focused on the Painted Bunting (*Passerina ciris*), which is a small Neotropical migrant songbird that breeds in

the southwestern part of the United States and along the coasts of South Carolina, Georgia, and Florida and winters in middle America (Thompson 1991a and 1991b). I implemented a NGS approach to generate thousands of sequences randomly distributed across the genome of the Painted Bunting. I then selected Painted Bunting sequences that were conserved across avian species and designed primers to amplify highly variable regions of DNA within coding regions in multiple species of migratory and non-migratory birds. In parallel, I built an extensive library of candidate migration genes based on annotations extracted from the Mouse Genome Informatics database (Fig 2). By crossing the two datasets, namely, the gene list obtained through NGS and the gene list obtained through phenotypic annotations, I identified novel candidate genes and tested their validity by using previously published data on phenotypic variation in migration among avian species. Data on genetic variation and migration speeds for a population of Painted Buntings in Oklahoma (US) that I used to test intra-species genotype/phenotype correlations came from Contina et al. (2013). Interspecific migration speed indexes were taken from La Sorte et al. (2013) to test for inter-species genotype/phenotype correlations.

Step 1. High-throughput DNA Sequencing

I implemented 454 technology (Roche) and performed a high-throughput DNA sequencing approach using a male Painted Bunting sample collected in Sinaloa (Mexico). I focused on 15,000 DNA fragments (about 30%) of the total pool of sequences generated with the 454 approach and used the Basic Local Alignment Search Tool (BLAST) to identify functional gene sequences conserved across avian species (Zhang et al. 2000). I optimized my search algorithm for highly similar sequences

(Megablast) and retained sequences at least 200bp in length and with 95% or higher identities with the model species, White-throated Sparrow (*Zonotrichia albicollis*).

Step 2. Mouse Genome Database

In the Mouse Genome Informatics database (http://www.informatics.jax.org/), I searched for the entire genotype/phenotype list, which returned 55,386 genotypes and 265,093 phenotypes as defined by annotations. To narrow down the list of genotypes for sequencing, I ranked the candidate genes by (a) assumed importance of the phenotype in relation to the trait "migration", and (b) base-pair identities with the avian genus *Gallus* and *Zonotrichia* for which extensive genomic data and detailed phenotypic annotations are available. In particular, I focused on four phenotype subgroups: the cardiovascular system, lipid metabolism, muscle growth and strength, and neuronal transmission (e.g. learning and memory). Then, I further refined our search by annotations selecting four subcategories of candidate genes involved in: "increased cardiac muscle contractility" (53 genes), "fat metabolism" (83 genes), "muscle growth" (60 genes), and "neuronal transmission" (40 genes). Next, I checked if the mammalian genes were orthologous in the chicken (*Gallus gallus*) and I conducted a BLAST search (blastn) to identify perfectly identical genomic regions (≥200 bp) and to define conserved regions (CRs) between the chicken (*Gallus sp*.) and the White-throated Sparrow (*Zonotrichia albicollis*), as a representative migrant and closely-related species of many songbirds (Passeriformes). Finally, I downloaded the gene sequences from NCBI (http://www.ncbi.nlm.nih.gov/) and designed the primers for tetra, penta, and esa nucleotide short tandem repeats (microsatellites) within the protein

coding regions and at multiple locations spread across the gene sequences using the software Primer3 (Rozen and Skaletsky, 1998).

Step 3. PCR and Cross-Species Validation

I initially tested primer pairs on two Painted Bunting DNA samples using 10 ul reactions containing at 2 ul of DNA, 1 ul of both the forward and reverse primer, 7 ul Qiagen Mastermix DNA polymerase. PCR was conducted at the following conditions: (1) 1 min at 95 1C; (2) 20 cycles of (a) 30 s at 95 1C, (b) 45 s at 65 1C at cycle 1 min—0.5 1C at every step, and (c) 60 s at 72 1C; (3) 20 cycles of (a) 30 s at 95 1C, (b) 45 s at 55 1C, and (c) 60 s at 72 1C; and (4) 5min at 72 1C. Because Painted Bunting amplicon sequences could also be unambiguously mapped back to White-throated Sparrow using BLAST, I am confident that the amplicons characterize the target genes. I then performed sequencing for each amplicon on a panel of 19 male Painted Bunting samples collected at the Wichita Wildlife National Refuge in Oklahoma (WWNR; N34.4 – W98.4), for which geologger data were also available, and on 21 species of migratory and non-migratory birds collected through MAPS banding programs and UCLA feather tissue collections (Table 1).

Step 4a. Merging the Datasets; Intra-Species Analysis

Based on the geologger analysis results for a breeding population of Painted Buntings in southwestern Oklahoma (WWNR), Contina et al. (2013) showed a great degree of variation in terms of duration of migration during the first part of the fall migratory movement between Oklahoma and Sinaloa (Mexico) for several male individuals ranging from a few days to almost a month. For a detailed report of the methods implemented for geologger tag deployment and data analysis, see Contina et

al. (2013). Here, I note that I obtained tissue samples for those 19 individuals fitted with geologgers and stored them at the Oklahoma Biological Survey, University of Oklahoma making it possible to genotype them at the novel candidate migration genes presented in this paper. Additionally, I ran a further genotyping effort that included twelve neutral but highly variable microsatellite sequences showing short-tandem-repeat polymorphisms. These sequences were previously developed for a separate forensic study of the Painted Bunting (Contina et al. in prep.) and were included as a negative control when we tested for correlation between duration of migration and candidate gene variation in the Oklahoma bunting population (Table 2). Finally, I calculated the Pearson correlation coefficient with a significant alpha level < 0.05 between allele size for each individual at these control sequences, at the novel candidate migration genes, and a variable phenotypic trait related with migratory behavior in the Oklahoma buntings. I defined this variable phenotypic trait as the time (days) necessary to fly from the breeding ground in Oklahoma to the molting ground in Sinaloa in late summer; these methods are described in detail elsewhere (Bridge et al. 2011; Contina et al. 2013; Contina et al. in prep).

Step 4b. Merging the Datasets; Inter-Species Analysis

I gathered published analyses to acquire detailed information for the phenotype "speed of migration" (km/day) over the entire annual cycle across different species of songbirds. I found estimates only for 13 out of the 21 species that I genotyped (La Sorte et al. 2013). Therefore, I limited my correlation analysis to these 13 species (Table 5; modified from La Sorte et al. 2013). Because the estimates of migration speed indexes calculated by LaSorte were at the population, rather than individual-level, I treated these

as an index of migration speed rather than attempting to convert them to units of individual movement rates. I calculated the Pearson correlation coefficient with a significant alpha level < 0.05 between allele size of each species for the novel candidate migration genes and a pool of phenotypic traits generally assumed to be linked with migratory behavior. For this inter-species analysis, I considered three migratory phenotypic traits: body mass, annual migration distance, and speed of migration (km/day). I also implemented general linear models (GLM) to investigate how much variation was accounted for by each variable in the explanatory model. Finally, I used three species (not included in the model), the Mourning Warbler (*Geothlypis philadelphia*), Purple Martin (*Progne subis*), and Painted Bunting for which population and individual migratory data were available, and tested the model predictions on these species. All the statistical analyses were computed in R and XLSTAT (R Development Core Team 2005; Addinsoft 2014).

Results

The 454 high-throughput DNA sequencing on a Painted Bunting sample yielded 48,694 sequences. These sequences ranged from 60bp to 510pb and provided an average of about 17 million nucleotides representing about 2% of the entire chicken genome (*Gallus gallus*). However, because I limited our analysis to 15,000 sequences out of those 48,694, only about 750 sequences undoubtedly aligned with avian protein coding regions (exons). Furthermore, after filtering by phenotype subcategories associated with candidate genes involved in increased cardiac muscle contractility, fat metabolism, muscle growth, and neuronal transmission, I identified a set of 87 Painted Bunting gene sequences conserved across multiple avian species, at least 200 bp in

length, and with 95% or higher identities with the model species White-throated Sparrow (Table 1S, suppl. material). I then clustered this set of genes by annotations to reduce phenotypic redundancy and designed primers for those genes that presented the highest degree of variation in terms of short sequence repeats (microsatellites) within the exons (Table 2S, suppl. material).

Thus, this further screening reduced my list of candidate genes down to twelve (Table 4S, suppl. material). After designing and testing the quality of the primers for these twelve genes, I only focused on four genes based on amplification success rate (due to species-specific primers) and allelic richness within the Painted Bunting population and across species of migratory birds. These four novel candidate genes were: ADRA1d; ANKRD17; CISH; and MYH7 (Table 4). For a short description of the gene functions see annotations in Table 4S in the supplementary material and discussion.

Results: Intra-Species Level

All four of the novel candidate genes showed little variation within the Oklahoma population of Painted Buntings (Table 2). ADRA1d and ANKRD17 showed only 2 alleles (N = 4) and were dismissed from further analysis. CISH showed 6 alleles (N = 10) but it was difficult to amplify across all the 19 individuals (amplification success rate 52%). MYH7 had an amplification success rate of 100%, but almost zero variation with only one individual showing a distinct allele with 1bp difference (Table 2). Two of the 12 neutral sequences (sequence ID# PABU-12 and sequence ID# PABU-42) varied in terms of number of microsatellite repeats that positively correlated with

the duration of migration for the first part of the fall migratory movement between Oklahoma and Sinaloa (data presented as averages in Fig 1a; 1b, and Table 3).

Results: Inter-Species Level

At the interspecific level, I found an interesting degree of variation. ADRA1d was amplified in 11 out of the 21 species tested, and showed three alleles ranging from 397bp to 399bp. ANKRD17 was amplified in 11 out of the 21 species tested and showed 6 alleles ranging from 277bp to 301bp. CISH was amplified in 10 out of the 21 species tested and showed 13 alleles ranging from 358bp to 411bp. MYH7 was amplified in all of the 21 species considered and revealed high base pair variation, ranging from 411bp to 473bp, and showed 13 different alleles (Table 2). Furthermore, the multiple regression model that best explained the data (Akaike's IC = 46.39) revealed that the speed of migration (km/day) was positively correlated with the average allele size of MYH7 (standardized regression coefficient = 0.7, $p = 0.004$) and negatively correlated with body mass across eleven avian species for which I had migration speed data (standardized regression coefficient = -0.4, $p = 0.048$). The model explained about 70% of the variance and was statistically significant (ANOVA: $F = 13.5$, $p < 0.003$). The results of the general linear models are presented in Table 3S (suppl. material).

Inter-Species Model Validation

Because I identified variables correlated with speed of migration indexes, I built a simple model equation that can be easily tested against other species. However, to calculate the migration speed estimates at the individual level, I introduced a correction

factor theta (θ) of 10x based on the assumption that the model is built upon population-level migration speed calculated on the species distribution centroid and this parameter severely reduces the actual individual migration speed. Thus, the speed of migration equation is described by the following terms:

$$SOM = (\theta)aX + (\theta)bY + c$$

where SOM is the speed of migration (km/day), X is the allele size of MYH7, Y is the body mass (grams) of the migrant, c is the intercept, a and b are the slopes, θ is the correction factor (constant) for computations at the individual level and is set at 10x.

La Sorte et al. (2013) estimated that the speed of migration for the Mourning Warbler is 59.7 km/day. This is a population-level migration speed estimate, and it is based on the great-circle distance measured sequentially between pairs of estimated population centroids. So it is important to remember that this is not the speed of a single individual. Our speed of migration equation predicts that the population-level migration speed for the Mourning Warbler should range between 22 km/day and 31 km/day.

Stutchbury et al. (2009) provided direct estimates for the speed of migration in the Purple Martin using geologgers. According to these estimates, a single Purple Martin has migration speed of 577 km/day. Our model equation based on genetic and body mass data predicts that a single Purple Martin could reach migration speed of 544 km/day (range 488-598 km/day), which is comparable to Stutchbury's et al. findings inferred with a direct method (geologgers). However, to calculate this speed estimate, I used a correction factor (θ) based on the assumption that my model is built upon population-level migration speed and not individual data.

Finally, Contina et al. (2013) published geologger data for several individuals of the Painted Bunting across multiple years. Because DNA samples were available for the same birds tagged with geologgers, I tested my speed equation on a specific individual that migrated from the breeding ground to the molting ground in about 9 days (PABU-08; band# 222150752). Considering a breeding-to-molting ground distance of 7000 km (OK-Sinaloa), I could estimate that this particular bird flew at a speed of 778 km/day. According to my equation, the Painted Bunting considered in this example should have a migration speed of 643 km/day.

Discussion

In this paper I present the results of an extensive candidate gene approach aimed to detect and validate potentially functional variations in novel candidate migration genes.

My results focus on the candidate migratory genes ADRA1d, ANKRD17, CISH, and MYH7. These genes code for different, but important, biological functions ranging from abnormal cardiovascular system physiology, enhanced coordination, increasing vascular permeability, increasing lymphocyte cell number, to decreasing body weight and enlarging skeletal muscle fiber size (for a complete list of annotations see Table 4S, suppl. material). Therefore, considering the importance of these physiological functions, I presume that even small allele mutations in the exons could alter gene expression and affect resulting phenotypes (Sutter et al. 2007) and I am puzzled by the degree of variation revealed at the inter-species level. In particular, ADRA1d did not show marked polymorphism, but ANKRD17, CISH, and MYH7 showed multiple alleles (6 to

15) and striking base pairs dissimilarities to a maximum of 53bp range difference in the case of CISH.

The prominent allele polymorphism at MYH7 between species is absent within a population of Painted Buntings known to show highly variable migratory behaviors in their duration of autumn migration (Table 2). A possible explanation for this discrepancy is that MYH7 could set the muscular physiological limits for each species (e.g. each species has its own MYH7 allele), but other ecological factors might influence the migratory behavior at the individual level. Therefore, each species might have its own "potential migratory muscle strength and migratory speed" that maintains individual variation between individuals and populations.

I discovered two uncharacterized polymorphic genes that clearly correlated with the duration of migration in the Painted Buntings for their first part of the migratory movement in late summer from the breeding ground in Oklahoma to the stop-over and molting site in Sinaloa (Bridge et al. 2011; Contina et al. 2013). The correlations of these two uncharacterized polymorphic genes need to be interpreted with caution, and further inter-population tests should be conducted to see if the pattern holds across the species range (Contina et al. in prep). Nevertheless, I note that previous studies have confirmed that microsatellites at the untranslated regions (UTR) of some *Arabidopsis* genes, for example, can influence DNA transcription in this species (Kashi and Soller, 1999). Similarly, the meaning of the correlation between the number of microsatellite repeats within the exon of MYH7 and the speed of migration (km/day) is difficult to explain with the current data. I speculate that the intron sequences containing the short tandem repeats is located in proximity of a promoter region adjacent to the

transcriptional start site, and therefore affects gene expression (Wang et al. 2013). I note that exon microsatellite sequences are linked to gene expression in several eukaryotic organisms when repeat clusters are located upstream of the promoter region (O'Donnell et al. 1994). Several studies have indisputably demonstrated that microsatellite mutations (e.g. variation in number of repeats) are not always neutral and may be even the root of several neurodegenerative disorders in humans. For example, Huntington's disease is triggered by simple proliferations of a CAG motif repeat occurring at a certain protein gene in humans (Moxon and Wills, 1999). My two novel uncharacterized polymorphic loci could be used as heuristic tools to predict duration of migration among different populations as well as the variation in number of microsatellite repeats within the exon of MYH7 and its correlation with the speed of migration (km/day) could be used for investigating phenotypic variation across different species.

There is a vast body of literature that addresses the link between body mass and speed of flight and migration (Pennycuick 1969 and references therein) yet, to the best of my knowledge, genetic data have never been included in any bird movement modeling attempts. My results are encouraging because I obtained comparable migration speed results when these were solely calculated with direct methods (e.g. geologgers), or when estimates were computed by simulation models on the basis of genetic data and body weight. My example of the Purple Martin is striking, given that a previously published migration speed estimate (km/day) for a single individual is only 6% off our average model estimate exclusively based on genetics and body mass.

Finally, I observe that even though the 454 high-throughput DNA sequencing reads generated large data files, and that my wide-ranging library assembly of candidate genes started at 55,386 genotypes and 265,093 phenotypes classified by annotations, I narrowed down the number of candidate migration genes to 12. This result should not be surprising because the purpose of crossing datasets is to reduce the number of candidate genes to a manageable number for a realistic sequencing effort. The identification of genes that are conserved across species, but that also show polymorphism usually results in elimination of many sequences from the search pool. For example, Steinmeyer et al. (2009) studied the clock genes and circadian behavior in the Blue Tit (*Cyanistes caeruleus*) and searched for tandem repeats in exon regions from an initial query of 206 and 438 genes extracted from the NCBI and the UCSC database, respectively, to only find five viable candidate genes. Similarly, Poelstra et al. (2013) implemented an extensive candidate gene approach to study pigmentation genes in the Eurasian crows (*Corvus* sp.) and sequenced three amplicons per gene in 37 pigmentation candidate genes starting from a pool of 142 genes extracted from the Mouse Genome Informatics database.

Conclusions

My results show the potential of interdisciplinary research in the field of avian migration and genetics. The candidate gene approach can now benefit from an unprecedented volume of gene annotations (e.g. list of functional gene associations to a particular phenotype/s) which are constantly updated and freely available. However, even though screening for novel candidate migration genes is now a facilitated task, it needs to be complemented with two more challenging steps consisting of rigorous

validation trials across populations and species. That is, the originality of my research is characterized by these two crucial, but often unrecognized validation steps: in my case the first validation step is through geologger migration data and DNA samples collected for a population of Painted Bunting in the US; the second validation step is at the inter-species level through migratory data over the full annual cycle gathered from the literature for different species of migratory birds. My results indicate that when these cross-validations are implemented, the final inferences and study conclusions can be significantly dissimilar across populations and species. Therefore, I urge that future studies need to systematically include cross-species validation that implement a candidate gene approach.

Acknowledgments

I thank the UCLA Center for Tropical Research, UCLA Institute of the Environment and Sustainability, the Institute for Bird Populations, the Monitoring Avian Productivity and Survivorship (MAPS) Program for collecting feather samples.

Figures and tables

Table 1. PABU genotypes at 42 and 12

Bird ID	Band ID	42-A	42-B	42-AV	12-A	12-B	12-AV	Migration (days)
PABU-07	222150751	153	163	158	162	178	170	15
PABU-05	222150752	163	167	165	151	162	156.5	7
PABU-15	222150754	162	186	174	171	182	176.5	14
PABU-06	222150755	163	179	171	171	183	177	9
PABU-20	222150757	152	163	157.5	151	162	156.5	1
PABU-01	222150758	153	163	158	151	163	157	2
PABU-09	222150760	164	175	169.5	166	178	172	5
PABU-16	251112762	163	167	165	170	179	174.5	15
PABU-14	251112838	163	171	167	170	178	174	10
PABU-21	251112852	163	175	169	166	171	168.5	16
PABU-19	251112854	163	183	173	162	178	170	13
PABU-13	251112858	163	175	169	174	174	174	7
PABU-02	251112859	163	167	165	140	150	145	2
PABU-11	251112861	167	175	171	162	171	166.5	10
PABU-22	251112864	175	179	177	166	178	172	29
PABU-10	251112865	163	175	169	150	171	160.5	5
PABU-04	251112866	163	175	169	151	162	156.5	9
PABU-12	251112883	152	163	157.5	167	171	169	9
PABU-17	251112890	152	163	157.5	158	171	164.5	5

Table 2. Candidate genes

PABU Band ID	ADRA1d-A	ADRA1d-B	ANKRD17-A	ANKRD17-B	CISH-A	CISH-B	MYH7-A	MYH7-B
222150751	398	398	278	298	386	392	413	469
222150752	n/a	n/a	n/a	n/a	387	393	413	469
222150754	n/a	n/a	n/a	n/a	385	393	413	469
222150755	n/a	n/a	n/a	n/a	393	393	413	469
222150757	n/a	n/a	n/a	n/a	385	385	413	469
222150758	n/a	n/a	n/a	n/a	393	393	413	469
222150760	399	399	277	298	385	393	414	469
251112762	398	398	277	298	393	393	413	469
251112838	n/a	n/a	n/a	n/a	0	0	413	469
251112852	n/a	n/a	n/a	n/a	0	0	413	469
251112854	n/a	n/a	n/a	n/a	0	0	413	469
251112858	n/a	n/a	n/a	n/a	0	0	413	469
251112859	n/a	n/a	n/a	n/a	0	0	413	469
251112861	n/a	n/a	n/a	n/a	0	0	413	469
251112864	n/a	n/a	n/a	n/a	0	0	413	469
251112865	n/a	n/a	n/a	n/a	0	0	413	469
251112890	398	398	278	298	385	394	413	469
222150719	n/a	n/a	n/a	n/a	385	393	413	469

Table 3. PABU correlation matrix

Correlation matrix (Pearson):

Variables	42	42b	42_AVA	DM61N-12	12-b	12_AVA	duration
42	**1**	**0.704**	**0.907**	0.220	0.231	0.232	**0.563**
42b	**0.704**	**1**	**0.937**	0.412	**0.502**	**0.470**	0.445
42_AVA	**0.907**	**0.937**	**1**	0.352	0.410	0.392	**0.540**
DM61N-12	0.220	0.412	0.352	**1**	**0.875**	**0.971**	**0.524**
12-b	0.231	**0.502**	0.410	**0.875**	**1**	**0.965**	**0.564**
12_AVA	0.232	**0.470**	0.392	**0.971**	**0.965**	**1**	**0.561**
duration	**0.563**	0.445	**0.540**	**0.524**	**0.564**	**0.561**	**1**

Values in bold are different from 0 with a significance level alpha=0.05

p-values:

Variables	42	42b	42_AVA	DM61N-12	12-b	12_AVA	duration
42	**0**	**0.001**	**0.000**	0.365	0.342	0.338	**0.012**
42b	**0.001**	**0**	**< 0.0001**	0.080	**0.028**	**0.042**	0.056
42_AVA	**< 0.0001**	**< 0.0001**	**0**	0.140	0.081	0.097	**0.017**
DM61N-12	0.365	0.080	0.140	**0**	**< 0.0001**	**< 0.0001**	**0.021**
12-b	0.342	**0.028**	0.081	**< 0.0001**	**0**	**< 0.0001**	**0.012**
12_AVA	0.338	**0.042**	0.097	**< 0.0001**	**< 0.0001**	**0**	**0.013**
duration	**0.012**	0.056	**0.017**	**0.021**	**0.012**	**0.013**	**0**

Values in bold are different from 0 with a significance level alpha=0.05

Figure a.1. Regression model for 42-AV.

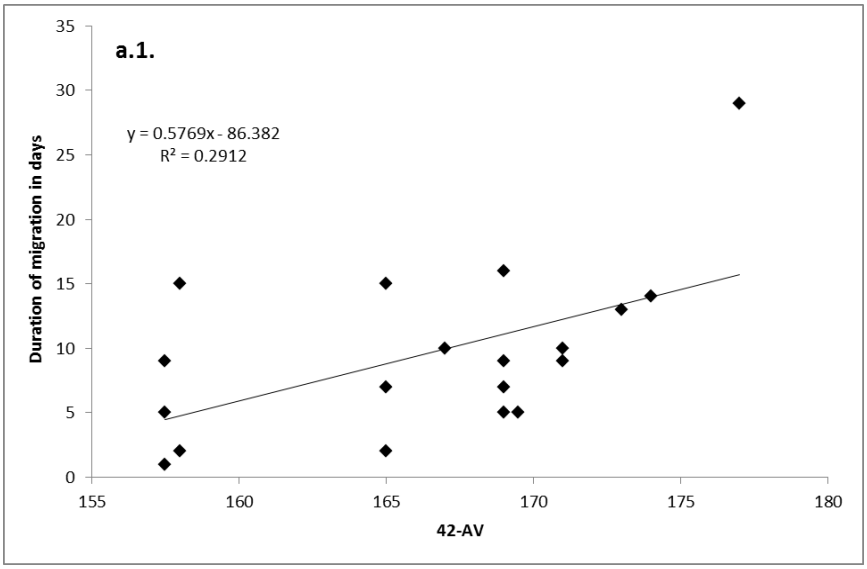


Figure b.2. Regression model for 12-AV

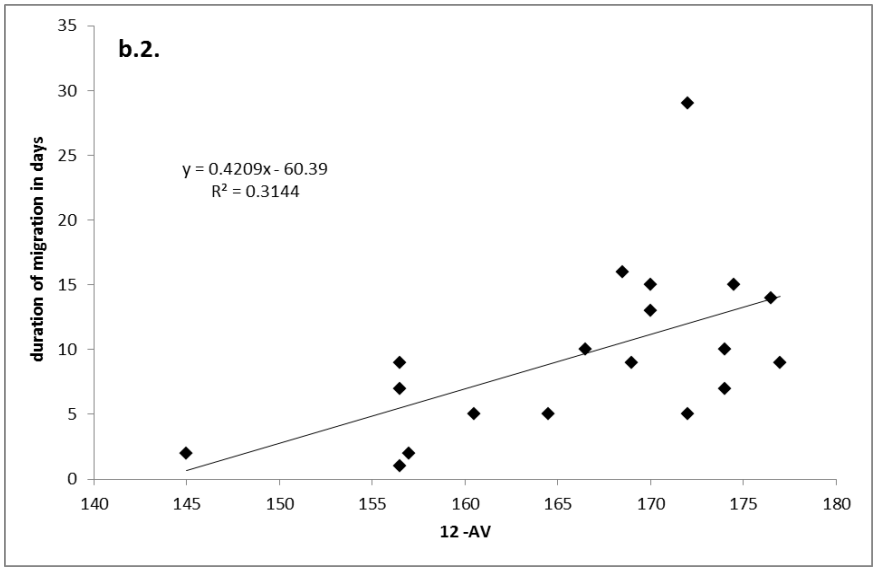
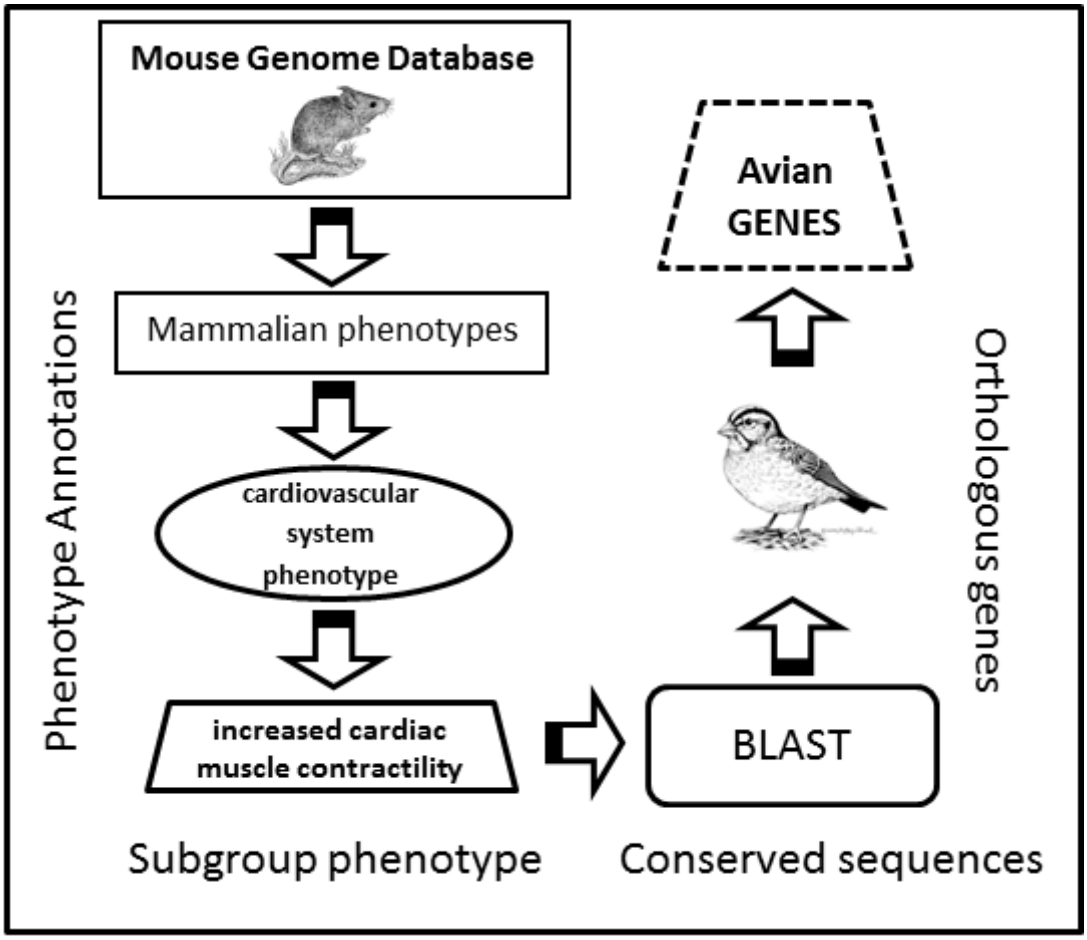


Table 4. Candidate gene primers

Gene	Primers	Primer seq.	Tm (°C)	Tm (°C)	Product Size (bsp)
ADRA1d	Forward Primer	TACTCCCTCAAGTACCCCACC	Tm (°C)	60.15	
Zoonotrichia albicollis	Reverse Primer	AGCTCCTGAAGTTGTGTCCCT	Tm (°C)	60.12	388
ANKRD18	Forward Primer	GAGCAGAGGATAAAAGAAGCCA	Tm (°C)	60.18	
Zoonotrichia albicollos	Reverse Primer	GGAGTCAGTAGTCCCAGATGCT	Tm (°C)	60.64	290
CISH	Forward Primer	TTAAAGTGCTGTTCCCTGACCT	Tm (°C)	60.165	
Zoonotrichia albicollos	Reverse Primer	AACCCAGAGCCAAAATCCTAAT	Tm (°C)	60.199	359
FMR1	Forward Primer	AGTTTGCCTTTGTGGAAGTGAT	Tm (°C)	60.04	
Zoonotrichia albicollos	Reverse Primer	TTTGAGCCACTCTTGTCTGCTA	Tm (°C)	60.19	388
MYH7	Forward Primer	CTGAAGAAGAAGATGGAGGGTG	Tm (°C)	60.239	
Zoonotrichia albicollos	Reverse Primer	TGGTGTTCTGTGAGTGGAGAAG	Tm (°C)	60.335	325
PIKFYVE	Forward Primer	CAATACCTGAAACATCCCCACT	Tm (°C)	60.11	
Zoonotrichia albicollos	Reverse Primer	CTTCTGCTCTCCAACTCCTTGT	Tm (°C)	60.053	388
INSR	Forward Primer	TGTCATCTCCATGTCCAACTCT	Tm (°C)	59.579	
Zoonotrichia albicollos	Reverse Primer	GATGTAAAACTTGACCAGCCCT	Tm (°C)	59.526	361
PRKDC	Forward Primer	ACAGTATGGAGGTGGATGGAGA	Tm (°C)	60.766	
Zoonotrichia albicollos	Reverse Primer	GGAAGGCTGTTATGTTCTTGCT	Tm (°C)	59.789	357
RYR2	Forward Primer	TCTGTTTGTTCCTCTCATCAAGC	Tm (°C)	60.787	
Zoonotrichia albicollos	Reverse Primer	ATCACTTCATTGTACCGGAAGC	Tm (°C)	60.381	390
NCOR1	Forward Primer	GATTGCTTGCTGGTCTTGATCT	Tm (°C)	60.772	
Zoonotrichia albicollos	Reverse Primer	CATACCCAGTTCTCCTTTCCTG	Tm (°C)	59.991	320
WWC2	Forward Primer	CTGCTTAGATTCATCATCCACG	Tm (°C)	59.73	
Zoonotrichia albicollos	Reverse Primer	TCCTTCATTATTCCTAGCTGGC	Tm (°C)	59.722	371
PTPRG	Forward Primer	AAGGATGAAAAGAGTGAAAGCG	Tm (°C)	59.892	
Zoonotrichia albicollos	Reverse Primer	TTCGGAGGAGGATTTCTTGTAG	Tm (°C)	59.72	368

Figure 1. The candidate gene approach.



References

Addinsoft. 2014. XLSTAT 2014, Data analysis and statistics software for Microsoft Excel, http://www.xlstat.com. Paris, France.

Åkesson, S. a. H., Anders. 2007. How Migrants Get There: Migratory Performance and Orientation. BioScience 57(2): 123-133.

Alberts B., Johnson A., Lewis J., Raff M., Roberts K.,Walter P. 2002. Molecular biology of the cell. 4th edn. New York: Garland Science.

Alerstam, T. 1990. Bird migration. Cambridge University Press, Cambridge.

Alerstam, T. and A. Hedenström. 1998. The Development of Bird Migration Theory. Journal of Avian Biology 29(4): 343-369.

Alerstam, T., Hedenström, A, Åkesson, S. 2003. Long-distance migration: evolution and determinants. Oikos 103(2): 247-260.

American Ornithologists' Union. 1957. Checklist of North American birds. 5th ed. Am. Ornithol.Union, Washington, D.C.

American Ornithologists' Union (AOU). 1998. Checklist of North American birds. 7th Edition. American Ornithologists' Union, Washington, D.C.

Arizmendi Arriaga, M. del C., 2006. Evaluación del estado de conservación de las poblaciones de *Passerina ciris, P. leclancherii, P. amoena, P. rositae, P. caerulea, Cardinalis cardinalis y C. sinuatus* en México: recomendaciones para su conservación y aprovechamiento sustentable. Universidad Nacional Autónoma de México. Facultad de Estudios Superiores Iztacala.

Askins, R. A. 1993. Population trends in grassland, shrubland, and forest birds in eastern North America. Pages 1–34 in Current Ornithology, vol. 11 (D. M. Power, Ed.). Plenum Press, New York.

Avise, J. C. 2004. "Chapter 7: Speciation and hybridization." In Molecular Markers, Natural History and Evolution (2nd edn.). Chapman and Hall, New York.

Avise, J.C. 2008. The history, purview, and future of conservation genetics. Pp. 5-15 in: Conservation Biology: Evolution in Action, S.P. Carroll and C.W. Fox (eds.). Oxford University Press, Oxford, U.K

Bai, M.-L., and D. Schmidt. 2012. Differential migration by age and sex in central European Ospreys *Pandion haliaetus*. Journal of Ornithology 153:75–84.

Bairlein F., Norris D.R., Nagel R., Bulte M., Voigt C.C., Hussel D., Fox J.W., Hussel D.J.T., Schmaljohann H. 2012. Cross-hemisphere migration of a 25-gram songbird. Biol Lett 8: published online. doi:10.1098/rsbl.2011.1223

Baker C.S. 2008. A truer measure of the market: the molecular ecology of fisheries and wildlife trade. Molecular Ecology. 17(18):3985-98

Baker, C.S. and S.R. Palumbi. 1994. Which whales are hunted? Molecular genetic evidence for illegal whaling. Science 265: 1538-1539.

Bauchinger, U. and Biebach, H. 2001. Differential catabolism of muscle protein in garden warblers (*Sylvia borin*): flight and leg muscle act as a protein source during long-distance migration. J. Comp. Physiol. B 171, 293-301.

Beaumont M.A., Zhang W., Balding D.J. 2002. Approximate Bayesian computation in population genetics. Genetics, 162, 2025–2035.

Ben-Shahar, Y. et al. 2002. Influence of gene action across different time scales on behaviour. Science 296, 741–744

Berthold P., Gwinner E., Sonnenschein E. (Eds.). 2003. Avian Migration. Springer-Verlag, Berlin.

Bischof, R., L. E. Loe, E. L. Meisingset, B. Zimmermann, B. Van Moorter, and A. Mysterud. 2012. A migratory northern ungulate in the pursuit of spring: Jumping or surfing the green wave? American Naturalist 180:407–424.

Blake, J.A., Bult, C.J., Eppig, J.T., Kadin, J.A., and Richardson, J.E. 2014.The Mouse Genome Database: integration of and access to knowledge about the laboratory mouse. Nucleic Acids Res. 42, D810–D817. doi:10.1093/nar/gkt1225

Boutin-Ganache I., Raposo M., Raymond M., Deschepper C.F. 2001. M13-tailed primers improve the reliability and usability of microsatellite analyses performed with two different allele sizing methods. Biotechniques 31 (1) 24-26.

Bowen, G. J., and J. Revenaugh. 2003. Interpolating the isotopic composition of modern meteoric precipitation. Water Resources Research 39:1299–1311.

Bowlin M. S., Bisson I. A., Shamoun-Baranes J., Reichard J. D., Sapir N., Marra P. P., Kunz T. H., Wilcove D. S., Hedenstrom A., Guglielmo C. G., Akesson S., Ramenofsky M., Wikelski M. 2010. Grand Challenges in Migration Biology. Integrative and Comparative Biology 50: 261-279.

Bridge, E. S., Fudickar, Adam M., Kelly, Jeffrey F., Contina, Andrea, Rohwer, Sievert. 2011. Causes of bimodal stable isotope signatures in the feathers of a molt-migrant songbird. Canadian Journal of Zoology 89(10): 951-959.

Bridge, ES, K Thorup, MS Bowlin, PB Chilson, RH Diehl, RW Fléron, P Hartl, R Kays, JF Kelly, WD Robinson, M Wikelski. 2011. Technology on the move: Recent and forthcoming innovations for tracking migratory birds. BioScience 61:689-697

Brohede .J, Primmer C.R., Moller A., Ellegren H. 2002. Heterogeneity in the rate and pattern of germline mutation at individual microsatellite loci. Nucleic Acids Res 30:1997–2003. doi:10.1093/nar/30.9.1997.

Brownstein M.J., Carpten J.D., Smith J.R. 1996. Modulation of non-templated nucleotide addition by tag DNA polymerase: primer modifications that facilitate genotyping. BioTechniques, 20, 1004 –1010.

Buler, J. J., F. R. Moore, et al. 2007. A Multi-Scale Examination of Stopover Habitat Use by Birds. Ecology 88(7): 1789-1802.

Butler, L. K., M. G. Donahue, and S. Rohwer. 2002. Molt-migration in Western Tanagers (*Piranga ludoviciana*): Age effects, aerodynamics, and conservation implications. Auk 119:1010–1023.

Carter, K. A., P. Rodriguez, and M. T. Dunn. 2008. An updated flora of the Wichita Mountains Wildlife Refuge. Oklahoma Native Plant Record 8:45–46.

Chilson, P.B., E. Bridge, W.F. Frick, J.W. Chapman and J.F. Kelly. 2012. Radar aeroecology:exploring the movements of aerial fauna through radio-waveremote sensing. Biology Letters . doi:10.1098/rsbl.2012.0384.

Cole, T.V.N., Hamilton, P., Henry, A.G., Duley, P., Pace III, R.M., White, B.N., Frasier, T. 2013. Evidence of a North Atlantic right whale Eubalaena glacialis mating ground. Endang Species Res 21:55−64

Contina, A., E.S. Bridge, N.S. Seavy, J.M. Duckles, and J.F. Kelly. 2013. Using geologgers to investigate bimodal isotope patterns in Painted Buntings (*Passerina ciris*). Auk 130:265-272.

Contina, A., E.S. Bridge, J.D. Ross, and J. F. Kelly. In prep. Testing candidate genes for migration in the Painted Bunting.

Contina, A., R.J. Harrigan, J.P. Pollinger, E.S. Bridge, J.F. Kelly, and T.B. Smith. In prep. A novel library of microsatellite loci isolated in the Painted Bunting.

Cornuet J, Ravigné V, Estoup A. 2010. Inference on population history and model checking using DNA sequence and microstaellite data with the software DIYABC (v.1). Bmc Bioinformatics 11: 401.

Craig, H. 1957. Isotopic standards for carbon and oxygen and correction factors for mass-spectrometric analysis of carbon dioxide. Geochimica et Cosmochimica Acta 12:133–149.

Craig, H. 1961. Standard for reporting concentrations of deuterium and oxygen-18 in natural waters. Science 133:1833–1834.

Cruz-Romo, L. y A. Oliveras de Ita. 2011. Conceptos ecológicos, métodos y técnicas para la conservación y aprovechamiento de aves canoras, de ornato y psitácidos. In: Sánchez, O. et al. (ed.) Temas sobre conservación de vertebrados silvestres en México. Semarnat e INE, Mexico. Pp. 121-148.

Csillery K, Blum MGB, Gaggiotti F, François O. 2010. Approximate Bayesian computation (ABC) in practice. Trends Ecol. Evol. 25:410–18.

Dalloul RA, Long JA, Zimin AV, Aslam L, Beal K, et al. 2010. Multi-Platform Next-Generation Sequencing of the Domestic Turkey (*Meleagris gallopavo*): Genome Assembly and Analysis. PLoS Biol 8(9): e1000475. doi:10.1371/journal.pbio.1000475

De Volo, S. B., R. T. Reynolds, M. R. Douglas, and M. F. Antolin. 2008. An improved extraction method to increase DNA yield from molted feathers. Condor 110: 762–766.

Deutschlander, M. E. and R. Muheim. 2009. Fuel reserves affect migratory orientation of thrushes and sparrows both before and after crossing an ecological barrier near their breeding grounds. Journal of Avian Biology 40(1): 85-89.

Earl D.A. 2009. Structure Harvester v 0.56.3 http://taylor0.biology.ucla.edu/struct_harvest/.

Elrod, M. L., N. E. Seavy, R. L. Cormier, and T. Gardali. 2011. Incidence of eccentric molt in first-year Wrentits increases with fledge date. Journal of Field Ornithology 82:325–332.

Enard, W., P. Khaitovich, J. Klose et al. 2002. Intra- and interspecific variation in primate gene expression patterns. Science 296:340–343.

Eskew, C. T. 1938. The flowering plants of the Wichita Mountains Wildlife Refuge. American Midland Naturalist 20:695–703.

Evanno G., Regnaut S., Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology 14: 2611–2620.

Excoffier L., Lischer H.E.L. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Molecular Ecology Resources 10: 564–567.

Foll, M., and O.E. Gaggiotti, 2006. Identifying the environmental factors that determine the genetic structure of Populations. Genetics 174 : 875-891.

Francois, O., and Durand, E. 2010. Spatially explicit Bayesian clustering models in population genetics. Molecular Ecology Resources, DOI: 10.1111/j.1755-0998.2010.02868.x.

Frankham, R., Ballou, J.D., Briscoe, D.A. 2002. Introduction to conservation genetics Cambridge University Press, Cambridge.

Fryxell, J. M., and T. Avgar. 2012. Animal migration: Catching the wave. Nature 490:182–183.

Gautier M. and Vitalis R. 2013. Inferring population histories using genome-wide allele frequency data. Molecular Biology and Evolution. 30(3):654–668.

González-Solís, J., M. Smyrli, T. Militão, D. Gremillet, T. Tveraa, R. A. Phillips, and T. Boulinier. 2011. Combining stable isotope analyses and geolocation to reveal kittiwake migration. Marine Ecology Progress Series 435:251–261.

Gosler A.G. 1991. On the use of greater covert moult and pectoral muscle as measures of condition in passerines with data for the Great Tit Parus major. Bird Study 38: 1–9. doi: 10.1080/00063659109477061

Goudet J. 2001. FSTAT, a program to estimate and test gene diversity and fixation indices (version 2.9.3). Available: http://www2.unil.ch/popgen/softwares/fstat.htm.

Goudet, J., M. Raymond, T. de Meeüs and F. Rousset. 1996. Testing differentiation in diploid populations. Genetics 144: 1933-1940.

Guglielmo, C. G., Haunerland, N. H., Hochachka, P. W. and Williams, T. D. 2002. Seasonal dynamics of flight muscle fatty acid binding protein and catabolic enzymes in a migratory shorebird. Am. J. Physiol. Regul. Integr. Comp. Physiol. 282, R1405-R1413.

Guo, S. and Thompson, E. 1992 Performing the exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics 48:361-372.

Hall, J.C. 1995. Tripping along the trail to the molecular mechanisms of biological clocks. Trends Neurosci. 18, 230–240.

Hamilton, R. A. 2001. Records of caged birds in Baja California. Pp. 254–257 in Monographs in Field Ornithology No. 3. American Birding Association, Colorado Springs, CO.

Helms, C. and W. Drury. 1960. Winter and migratory weight and fat studies on some North American buntings. Bird Banding 31: 1-40

Herr, C., P. Sykes, and J. Klicka. 2011. Phylogeography of a vanishing North American songbird: the Painted Bunting (*Passerina ciris*). Conservation Genetics 12(6): 1395-1410.

Herzog E.D., Grace M.S., Harrer C., Williamson J., Shinohara K., Block G.D. 2000. The role of Clock in the developmental expression of neuropeptides in the suprachiasmatic nucleus. J Comp Neurol 424: 86–98.

Hobson, K. A., S. L. Van Wilgenburg, K. Larson, and L. I. Wassenaar. 2009. A feather hydrogen isoscape for Mexico. Journal of Geochemical Exploration 102:167–174.

Hobson, K. A., S. L. Van Wilgenburg, T. Piersma, and L. I. Wassenaar. 2012. Solving a migration riddle using isoscapes: House Martins from a Dutch village winter over West Africa. PLoS ONE 7(9):e45005.

Hoffmann, A.A. 1994. Behaviour genetics and evolution. In: Slater, P.J.B. & Halliday, T.R. (eds) Behavioural evolution. Cambridge: Cambridge University Press: 7–42.

Hosoya M., Kimura C., Ogi K., Ohkubo S., Miyamoto Y., Kugoh H., Shimizu M., Onda H., Oshimura M., Arimura A., Fujino M.1992. Structure of the human pituitary adenylate cyclase activating polypeptide (PACAP) gene. Biochimica et Biophysica Acta 1129:199-206.

Howell, S.N.G., and S. Webb. 1995. A guide to the Birds of México and Northern Central America. Oxford Univ. Press, New York.

Hubisz, M., Falush, D., Stephens, M., Pritchard, J. 2009. Inferring weak population structure with the assistance of sample group information. Molecular Ecology Resources, 9, 1322–1332.

Huntington's Disease Collaborative Research Group. 1993. A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. Cell: 72: 971–983.

Iñigo-Elias, E. E. 1986. Active trade threatens Mexican avifauna. TRAFFIC (U.S.A.) Vol. 6 (4): 6-7.

Iñigo-Elias, E. E., K. V. Rosenberg, and J. V. Wells. 2002. The Danger of Beauty. Birdscope, newsletter Summer 2002 Vol 16, no 3, of the Cornell Lab of Ornithology, Summer 2002. http://birds.cornell.edu/publications/birdscope.

Jahn, A. E., V. R. Cueto, J. W. Fox, M. S. Husak, D. H. Kim, D. V. Landoll, J. P. Ledezma, H. K. LePage, D. J. Levey, M. T. Murphy, and R. B. Renfrew. 2013. Migration timing and wintering areas of three species of flycatchers (*Tyrannus*) breeding in the Great Plains of North America. Auk 130:

Jakobsson M., Rosenberg N.A. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23: 1801–1806.

Jenkins, K. D., and D. A. Cristol. 2002. Evidence of differential migration by sex in White-throated Sparrows (*Zonotrichia albicollis*). Auk 119:539–543.

Johnsen A., Fidler A.E., Kuhn S., et al. 2007. Avian Clock gene polymorphism: evidence for a latitudinal cline in allele frequencies. Mol Ecol. 2007; 16(22): 4867–4880.

Kalinowski, S.T. 2011. The computer program STRUCTURE does not reliably identify the main genetic clusters within species: simulations and implication for human population structure. Heredity, 106, 625–632; doi:10.1038/hdy.2010.95.

Kashi Y., Soller M. 1999. Functional roles of microsatellites and minisatellites. In: Microsatellites: Evolution and Applications (eds Goldstein DB, Schlötterer C), pp. 10-23. Oxford University Press, Oxford.

Kelly, J. F., E. S. Bridge, A. M. Fudickar, and L. I. Wassenaar. 2009. A test of comparative equilibration for determining non-exchangeable stable hydrogen isotope values in complex organic materials. Rapid Communications in Mass Spectrometry 23:2316–2320.

Keyser, A. J., and G. E. Hill. 2000. Structurally based plumage coloration is an honest signal of quality in male Blue Grosbeaks. Behavioral Ecology 11:202–209.

Kitano, J., Ishikawa, A., Kume, M. and Mori, S. 2012. Physiological and genetic basis for variation in migratory behavior in the three-spined stickleback, *Gasterosteus aculeatus*. Ichthyol. Res.,59: 293–303.

Kwon J.M., Goate A.M. 2000. The candidate gene approach. Alcohol Res Health 24:164–8

La Sorte, F. A., D. Fink, W. M. Hochachka, J. P. DeLong and S. Kelling. 2013. Population-level scaling of avian migration speed with body size and migration distance for powered fliers. Ecology 94(8): 1839-1847.

Lande, R. 1998. Genetics and demography in biological conservation. Science, 241: pp. 1455–1460.

Liao, B. Y. & Zhang, J. 2006. Evolutionary conservation of expression profiles between human and mouse orthologous genes. Mol. Biol. Evol. 23, 530–540.

Liedvogel M., Åkesson S., Bensch S. 2011.The genetics of migration on the move. Trends in Ecology and Evolution, 26, 561–569.

Lisovski, S., and S. Hahn. 2012. GeoLight—Processing and analysing light-based geolocator data in R. Methods in Ecology and Evolution 3:1055–1059.

Lopez Medellin, X., and E. E. Iñigo-Elias. 2002. The Passeriformes Cage Trade in Mexico: a Challenge for Bird Conservation. Paper presented at The Third North American Ornithological Conference. New Orleans, Louisiana, Sept. 24-28, 2002.

Loria, D. E., and F. R. Moore. 1990. Energy demands of migration on Red-eyed Vireos, *Vireo olivaceus*. Behavioral Ecology 1(1): 24-35.

Lowther, P. E., S. M. Lanyon, and C. W. Thompson. 1999. Painted Bunting (*Passerina ciris*). In Birds of North America Online (A. Poole, Ed.). Cornell Lab of Ornithology, Ithaca, New York. Available at bna.birds.cornell.edu/bna/species/398.

Marra, P. P., and R. L. Holberton. 1998. Corticosterone levels as indicators of habitat quality: Effects of habitat segregation in a migratory bird during the non-breeding season. Oecologia 116:284–292.

Marra, P. P., K. A. Hobson, and R. T. Holmes. 1998. Linking winter and summer events in a migratory bird using stable-carbon isotopes. Science 282:1884–1886.

Martin, A. C., H. S. Zim, and A. L. Nelson. 1951. American Wildlife and Plants: A Guide to Wildlife Food Habits. McGraw-Hill, New York.

McFarlan, J. T., Bonen, Arend, Guglielmo, Christopher G. 2009. Seasonal up regulation of fatty acid transporters in flight muscles of migratory white-throated sparrows (*Zonotrichia albicollis*). Journal of Experimental Biology 212(18): 2934-2940.

Merlin C., Gegear R.J., Reppert S.M. 2009. Antennal circadian clocks coordinate sun compass orientation in migratory monarch butterflies. Science; 325(5948): 1700–1704.

Mlodinow, S. G. and R. A. Hamilton. 2005. Vagrancy of Painted Bunting (*Passerina ciris*) in the United States, Canada, and Bermuda. North American Birds 59:172–183.

Møller, A. P., E. Flensted-Jensen, et al. 2010. Climate change affects the duration of the reproductive season in birds. Journal of Animal Ecology 79(4): 777-784.

Moritz C.1994. Defining ‘evolutionarily significant units’ for conservation. Trends Ecol Evol 9:373–375

Mueller J.C., Pulido F., Kempenaers B. 2011. Identification of a gene associated with avian migratory behaviour. Proc Biol Sci.; 278(1719): 2848–2856.

Naef-Daenzer, B. 2007. An allometric function to fit leg-loop harnesses to terrestrial birds. Journal of Avian Biology 38: 404–407.

National Audubon Society (NAS). 2004a. Audubon Watchlist Species, Painted Bunting Passerina ciris. http://audubon2.org/webapp/watchlist/viewSpecies.jsp?id=156

National Audubon Society (NAS). 2004b. The Christmas Bird Counts. National Audubon Society. http://www.audubon.org/bird/cbc/

Norris, D. R., and P. P. Marra. 2007. Seasonal interactions, habitat quality, and population dynamics in migratory birds. Condor 109:535–547.

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*). Mol Ecol. 2007; 16(23): 4930–4941.

Ogden R., Dawnay N., McEwing R. 2009. Wildlife DNA forensics - bridging the gap between conservation genetics and law enforcement. Endangered Species Research, 9: 179-195.

Osborne, K.A. et al. 1997. Natural behaviour polymorphism due to a cGMP-dependent protein kinase of Drosophila. Science 277, 834–836

Paritte, J. M., and J. F. Kelly. 2009. Effect of cleaning regime on stable-isotope ratios of feathers in Japanese Quail (*Coturnix japonica*). Auk 126:165–174.

Paynter, Jr., R. A. 1970. Subfamily Cardinalinae. Pages 216-245 in Check-list of birds of the world.Vol. 13 (Paynter, Jr., R. A., Ed.) Mus. Comp. Zool. Cambridge, MA.

Pertea M., Salzberg S.L. 2010. Between a chicken and a grape: estimating the number of human genes. Genome Biol. 2010;11:206.

Peterson M.P., Abolins-Abols M., Atwell J.W. et al. 2013. Variation in candidate genes CLOCK and ADCYAP1 does not consistently predict differences in migratory behavior in the songbird genus Junco [v1; ref status: indexed, http://f1000r.es/11p] F1000Research 2013, 2:115 (doi: 10.12688/f1000research.2-115.v1)

Piersma, T., Gudmundsson, G. A. and Lilliendahl, K. 1999. Rapid changes in the size of different functional organ and muscle groups during refueling in a longdistance migrating shorebird. Physiol. Biochem. Zool. 72, 405-415.

Pittendrigh, C.S. 1993. Temporal organization: reflections of a Darwinian lock-watcher. Annu. Rev. Physiol. 55, 17–54.

Poelstra J.W., Ellegren H., Wolf J.B.W. 2013. An extensive candidate gene approach to speciation: diversity, divergence and linkage disequilibrium in candidate

pigmentation genes across the European crow hybrid zone. Heredity, 111, 467–473.

Pritchard J.K., Stephens M., Donnelly P. 2000. Inference of population structure using multilocus genotype data. Genetics 155: 945–959.

Pulido F., and Coppack T. 2004. Correlation between timing of juvenile moult and onset of migration in the blackcap (*Sylvia atricapilla*). Animal Behaviour 68:167-173.

Pulido, F. 2007. The genetics and evolution of avian migration. Biosciences, 57, 165–174.

Pulido, F., Berthold, P. and van Noordwijk, A.J. 1996. Frequency of migrants and migratory activity are genetically correlated in a bird population: evolutionary implications. Proc. Nat. Acad. Sci. USA 93: 14642–14647.

Pyle, P. 1997. Identification guide to North American birds. Part I. *Columbidae* to *Ploceidae*. Slate Creek Press, Bolinas, California.

Pyle, P., W. A. Leitner, L. Lozano-Angulo, F. Avilez-Teran, H. Swanson, E. Gómez-Limón, and M. K. Chambers. 2009. Temporal, spatial, and annual variation in the occurrence of molt-migrant passerines in the Mexican monsoon region. Condor 111:583–590.

R Development Core Team A Language and Environment for Statistical. R Foundation for Statistical Computing. Vienna, Austria. 2010.

Ramos, M. 1982. El comercio y la explotación de las aves vivas en México. INIREB, Cuadernos de Divulgación No. 8. Xalapa, Veracruz, México.

Rappole, J. H., and A. R. Tipton. 1991. New harness design for attachment of radio transmitters to small passerines. Journal of Field Ornithology 62:335–337.

Rappole, J. H., and D. W. Warner. 1980. Ecological aspects of migrant bird behavior in Veracruz, Mexico. Pp 353-393 In A. Keast and E. S. Morton (Eds.), Migrant birds in the Neotropics: Ecology, behavior, distribution, and conservation. Smithsonian Inst. Press, Washington, D.C.

Raymond M., and Rousset F. 1995. Genepop version 1.2: population genetics software for exact tests and ecumenicism. Journal of Heredity, 86, 248,249.

Raymond M. and F. Rousset. 1995 An exact test for population differentiation. Evolution 49:1280-1283.

Rich, T. D., C. J. Beardmore, H. Berlanga, P. J. Blancher, M. S. W. Bradstreet, G. S. Butcher, D. Demarest, E. H. Dunn, W. C. Hunter, E. Iñigo-Elias, J. A. Kennedy, A. Martell, A. Panjabi, D. N., Pashley, K. V. Rosenberg, C. Rustay, S. Wendt

and T. Will. 2004. Partners in Flight North American Landbird Conservation Plan. Cornell Lab of Ornithololgy. Ithaca, NY.

Riede, K. 2001. The Global Register of Migratory Species (GROMS): Present Status and Perspectives In: Riede, Klaus (Ed.): New Perspectives for Monitoring Migratory Animals - Improving Knowledge for Conservation. - Münster (Landwirtschafts-verlag), pp. 33 – 42.

Riede, K. 2004. The Global Register of Migratory Species - first results of global GIS Analysis. In: Werner, Dietrich (ed.): Biological Resources and Migration. Springer-Verlag, Berlin, pp. 211 – 218.

Robinson W. D., Bowlin M. S., Bisson I. A., Shamoun-Baranes J., Thorup K., Diehl R. H., Kunz T. H., Mabey S. E., Winkler D. W. 2010. Integrating concepts and technologies to advance the study of bird migration. Frontiers in Ecology and the Environment 8: 354-361.

Rohwer, S., and D. E. Irwin. 2011. Molt, orientation, and avian speciation. Auk 128:419–425.

Rohwer, S., K. A. Hobson, and V. G. Rohwer. 2009a. Migratory double breeding in Neotropical migrant birds. Proceedings of the National Academy of Sciences USA 106:19050–19055.

Rohwer, V. G., S. Rohwer, and M. F. Ortiz-Ramirez. 2009b. Molt biology of resident and migrant birds of the monsoon region of west Mexico. Ornitologia Neotropical 20:565–584.

Rohwer, S., L. K. Butler, and D. R. Froehlich. 2005. Ecology and demography of east–west differences in molt scheduling of Neotropical migrant passerines."Pages 87-105 in Birds of Two Worlds: The Ecology and Evolution of Migratory Birds (R.Greenberg and P. P. Marra, Eds.). Johns Hopkins University Press, Baltimore, Maryland.

Rosenberg N.A. 2004. DISTRUCT: a program for the graphical display of population structure. Molecular Ecology Notes 4: 137–138.

Rozen S., Skaletsky H. 1998. Primer3. Code available at: http://www-genome wi.mit.edu/genome_software/other/primer3.html. Accessed at: http://biotools.umassmed.edu/bioapps/primer3_www.cgi

Rundel, C.W., M.B. Wunder, A.H. Alvarado, K.C. Ruegg, R. Harrigan, A. Schuh, J.F. Kelly, R.B. Siegel, D.F. DeSante, T.B. Smith, J. Novembre. 2013. Novel statistical methods for integrating genetic and stable isotopic data to infer individual-level migratory connectivity. Molecular Ecology doi: 10.1111/mec.12393.

Saitou N. and Nei M. 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. Molecular Biology and Evolution 4:406-425

Sauer, J. R., J. E. Hines, and J. Fallon. 2003. North American Breeding Bird Survey Trend Results (Survey wide). The North American Breeding Bird Survey, Results and Analysis 1966 - 2002. Version 2003.1, USGS Patuxent Wildlife Research Center, Laurel, MD.

Sauer, J. R., J. E. Hines, G. Gough, I. Thomas, and B. G. Peterjohn. 1997. The North American Breeding Bird Survey Results and Analysis. Version 96.4. Patuxent Wildlife Research Center, Laurel, MD. [Online.] http://www.mbr-pwrc.usgs.gov/bbs/bbs96.html.

Schwartz, M. K., Luikart, G. and Waples, R. S. 2007. Genetic monitoring as a promising tool for conservation and management. Trends Ecol. Evol. 22, 25–33.

Seavy, N. E., D. L. Humple, R. L. Cormier, and T. Gardali. 2012. Establishing the breeding provenance of a temperate wintering North American passerine, the Golden-crowned Sparrow, using light-level geolocation. PloS ONE 7(4):e34886.

Secretaria de Medio Ambiente y Recursos Naturales (SEMARNAT). 2013. Aprovechamientos. <http://www.semarnat.gob.mx/temas/gestion-ambiental/vida-silvestre/aprovechamientos> consulted 10-december-2014.

Shipley, J.R., A. Contina, N. Batbayar, E.S. Bridge, A.T. Peterson, and J.F. Kelly. 2013. Niche conservatism and disjunct populations: a case study with Painted Buntings (*Passerina ciris*). Auk 130:476-486.

Sibley, C.G., and B.L., Monroe. 1990. Distribution and Taxonomy of Birds of the World. Yale University Press. New Haven, Connecticut.

Sprunt, A. Jr. 1954. Florida Bird Life. Coward-McCann, Inc. and National Audubon Society, New York.

Stach, R., S. Jakobsson, C. Kullberg, and T. Fransson. 2012. Geolocators reveal three consecutive wintering areas in the Thrush Nightingale. Animal Migration 1: 1–7.

Steinmeyer, C., Mueller, J. C., & Kempenaers, B. 2009. Search for informative polymorphisms in candidate genes: clock genes and circadian behaviour in blue tits. Genetica, 136(1), 109–117. doi:10.1007/s10709-008-9318-y

Storer, R. W.1951. Variation in the Painted Bunting (*Passerina ciris*), with special reference to wintering populations. Occ. Papers, Mus. Zool., Univ. Michigan, 532: 1–12.

Stutchbury, B. J. M., S. A. Tarof, T. Done, E. Gow, P. M. Kramer, J. Tautin, J. W. Fox, and V. Afanasyev. 2009. Tracking long-distance songbird migration by using geolocators. Science 323:896.

Sumner, M. D, and S. J. Wotherspoon. 2010. tripEstimation: Metropolis sampler and supporting functions for estimating animal movement from archival tags and satellite fixes. R package, version 0.0-33. Available online at cran.r-project.org/web/packages/tripEstimation.

Sumner, M. D., S. J. Wotherspoon, and M. A. Hindell. 2009. Bayesian estimation of animal movement from archival and satellite tags. PLoS ONE 4:e7324.

Sutter N.B., Bustamante C.D., Chase K., et al. 2007. A single IGF1 allele is a major determinant of small size in dogs. Science. 316:112–115.

Sykes, P. W., Jr., and S. Holzman. 2005. Current range of the eastern population of Painted Bunting (*Passerina ciris*). Part 1: Breeding. North American Birds 59: 4-17.

Sykes, P. W., Jr., S. Holzman, and E. E. Iñigo-Elias. 2007. Current range of the eastern population of Painted Bunting (*Passerina ciris*). Part II: Winter range. North American Birds 61:378–406.

Thompson, C. W. 1991a. Is the Painted Bunting actually 2 species - Problems determining species limits between allopatric populations. The Condor 93:987-1000.

Thompson, C. W. 1991b. The sequence of molts and plumages in Painted Buntings and implications for theories of delayed plumage maturation. Condor 93:209–235.

Thompson, C. W. 1992. A key for aging and sexing Painted Buntings. J. Field Ornithol. 63:445-454.

U.S. Fish and Wildlife Service (USFWS). 2008. Birds of Conservation Concern. U.S. Department of Interior, Fish and Wildlife Service, Division of Migratory Bird Management, Arlington, VA. Available online at http://www.fws.gov/migratorybirds/.

USFWS (U. S. Fish and Wildlife Service). 2004. North Carolina Endangered Species Home Page. Available online at http://nc-es.fws.gov/es/cntylist/brunswick.html.

Vaudry D., Falluel-Morel A, Bourgault S., et al.: Pituitary Adenylate Cyclase-Activating Polypeptide and Its Receptors: 20 Years after the Discovery. 2009. Pharmacol Rev; 61(3): 283–357.

Vaysse A., Ratnakumar A., Derrien T., Axelsson E., Rosengren Pielberg G., et al. 2011. Identification of genomic regions associated with phenotypic variation between dog breeds using selection mapping. PLoS Genet 7: e1002316.

Viau, A.E., Gajewski, K., Sawada, M.C. and Fines, P. 2006. Millennial-scale temperature variations in North America during the Holocene. Journal of Geophysical Research, 111: D09102, doi:10.1029/2005JD006031.

Voelker, G., and S. Rohwer. 1998. Contrasts in scheduling of molt and migration in Eastern and Western warbling-vireos. Auk 115:142–155.

Wang T, Zang W, Ma Y, Li M, Xuan X, Wang N, et al. 2013. DNA polymerase beta promoter mutations affect gene transcription, translation and the sensitivity of esophageal cancer cells to cisplatin treatment. Mol Biol Rep. 40(2):1333–9.

Webster, P. J., Holland, G. J., Curry, J. A., Chang, H.-R. 2005. Changes in Tropical Cyclone Number, Duration, and Intensity in a Warming Environment. Science 309(5742): 1844-1846.

Weir, B.S. and Cockerham, C.C. 1984. Estimating F-statistics for the analysis of population structure. Evolution 38:1358-1370.

Wijnstekers, W. 2011. The Evolution of CITES - 9th edition. International Council for Game and Wildlife Conservation.

Wilkening S., Chen B., Bermejo J.L., Canzian F. 2009. Is there still a need for candidate gene approaches in the era of genome-wide association studies? Genomics; 93:415–9.

Wiltschko, R. and Wiltschko, W. 2003. Orientation behavior of homing pigeons at the Gernsheim anomaly. Behavioral Ecology and Sociobiology 54(6): 562-572.

Winger, B. M., Lovette, Irby J., Winkler, David W. 2011. Ancestry and evolution of seasonal migration in the Parulidae. Proceedings of the Royal Society B: Biological Sciences.

Wong S.L., Zhang L.V., Roth F.P. 2005. Discovering functional relationships: biochemistry versus genetics. Trends Genet. 21(8):424–427. doi: 10.1016/j.tig.2005.06.006.

Wunder, M. B., C. L. Kester, F. L. Knopf, and R. O. Rye. 2005. A test of geographic assignment using isotope tracers in feathers of known origin. Oecologia 144:607–617.

Yong, W., D. M. Finch, F. R. Moore, and J. F. Kelly. 1998. Stopover ecology and habitat use of migratory Wilson's Warblers. The Auk 115: 829-842.

Yuri, T., and S. Rohwer. 1997. Molt and migration in the Northern Rough-winged Swallow. Auk 114:249–262.

Zatoń-Dobrowolska, M., Mucha, A., Wierzbicki, H., Morrice, D., Moska, M., Dobrowolski, M., & Przysiecki, P. 2014. Microsatellite polymorphism and its

association with body weight and selected morphometrics of farm red fox (Vulpes vulpes L.). Journal of applied genetics, 55(4), 475-484. 10.1007/s13353-014-0217-x

Zhang Z., Schwartz S., Wagner L., and Miller W. 2000. A greedy algorithm for aligning DNA sequences. Journal of Computational Biology, 7, 203-214. doi:10.1089/10665270050081478