MIGRATORY DIVIDES AND THE GENETIC BASIS OF REPRODUCTIVE ISOLATION

by

Kira Delmore

B.Sc.H. Queen's University, 2007

M.A., University of Calgary, 2009

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Zoology)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

July 2015

Abstract

Differences in seasonal migratory behaviour could be an important driver of ecological speciation. Many divergent groups form migratory divides on their breeding grounds; they breed adjacent to one another but use different routes to navigate around unsuitable areas on migration. Hybrids in divides are predicted to employ intermediate and inferior routes. I used light-level geolocators to track birds from the edges and center of a hybrid zone between inland and coastal Swainson's thrushes (*Catharus ustulatus*) in western Canada. These data provided the first direct identification of a migratory divide (Chapter 2) and support for the prediction that hybrids in divides take intermediate routes (Chapter 3). Hybrid routes crossed arid and mountainous regions, further suggesting that these routes are inferior.

I extended this work to examine the genetic basis of reproductive isolation between thrushes, assembling a reference genome and generating whole-genome sequence data for populations adjacent to the hybrid zone between these groups (Chapter 4). I documented genome-wide heterogeneity in genetic differentiation and uncovered patterns suggesting selective sweeps and variation in recombination generated this heterogeneity; within-population variation and absolute genetic differentiation were lower in regions of high relative differentiation and these reductions often coincided with centromeres and the Z chromosome. Genes associated with migration were concentrated in highly differentiated areas, further supporting migration's role in reproductive isolation between thrushes.

I complimented this work using a comparative approach to determine if patterns in the Swainson's thrush could be extended to other species (Chapter 5). Specifically, I compared rates of phenotypic divergence between sister pairs that form divides and those that do not. I considered phenotypic divergence a proxy for reproductive isolation and contrary to

expectations, found divergence was greater among taxa that do not form divides. This pattern could be explained by differential fusion, with sister pairs that do not form divides fusing into a single unit during periods of secondary contact unless they were sufficiently diverged phenotypically. Differences in migration would have permitted the persistence of pairs that form divides even without phenotypic differentiation. Under this scenario, migration serves as one of the major sources of speciation in North American birds.

Preface

A version of Chapter 2 has been published: **Delmore KE**, Fox JW, Irwin DE. 2012. Dramatic intraspecific differences in migratory routes, stopover and wintering sites revealed using light-level geolocators. *Proceedings of the Royal Society B: Biological Sciences* 279: 4582-4589. I designed, collected field data, conducted statistical analyses and wrote the paper. James Fox sold us geolocators at cost and provided advice during the analysis of data from these devices. Darren Irwin helped design this work and contributed ideas during the write-up phase.

A version of Chapter 3 has been published: **Delmore KE**, Irwin DE. 2014. Hybrid songbirds employ intermediate routes in a migratory divide. *Ecology Letters*, 17:1211. I designed the study, collected field data, conducted statistical analyses and wrote the paper. Darren Irwin helped design this work and contributed ideas during the write-up phase.

A version of Chapter 4 has been published: **Delmore KE**, Hübner S, Kane NC, Schuster R, Andrew RL, Câmara F, Guigo Roderic, Irwin DE. 2015. Genomic analysis of a migratory divide reveals candidate genes for migration and implicates selective sweeps in generating islands of differentiation. *Molecular Ecology*, 24:1873. I designed, collected field data, conducted statistical analyses and wrote the paper. Sariel Hübner, Nolan Kane, Richard Schuster, Rose Andew, Francisco Câmara and Guigo Roderic provided statistical advice. Darren Irwin helped design this work and contributed ideas during the write-up phase.

Work for chapter 5 was conducted in collaboration with Haley Kenyon, Ryan Germain and Darren Irwin. I designed, collected data, conducted statistical analyses and wrote the paper. Haley Kenyon helped collect data, conduct statistical analyses and contributed ideas during the

write-up phase. Ryan Germain provided statistical advice. Darren Irwin helped design this work and contributed ideas during the write-up phase.

All protocols involving the use of animals in this dissertation were approved by the UBC animal care committee (A09 – 0131, Geographic variation in birds of western Canada). Permits for this work were obtained from Environment Canada (master banding permit no. 10746 and scientific permit no. BC-10-0035). This work was supported by grants from the Faculty of Grad Studies at UBC, Zoology Department at UBC, Queen's University's Alumnae Association, Science Horizons Community Funding Program with Environment Canada, Canadian Wildlife Service, Sigma Xi, Hesse Research Fund at UBC, American Museum of Natural History, National Geographic, Society for the Study of Evolution, American Ornithologists Union, Animal Behavior Society, Society of Canadian Ornithologists and Wilson Ornithological Society.

Table of Contents

List of Tables

List of Figures

Acknowledgements

First and foremost, I would like to thank my supervisor Darren Irwin for his encouragement, advice and support. We faced quite a few challenges during my PhD and I never doubted that I could count on your support through each of them. I learned a lot from working with you and will benefit from this knowledge throughout my career. Along the same lines, I am grateful for the support I received from my supervisory committee: Kathy Martin, Loren Rieseberg and Dolph Schluter. I am especially thankful to Dolph, who accepted me into his lab when Darren was on sabbatical and (perhaps to his chagrin) failed to shake me from his lab after that point. Katie Lotterhos and Mike Whitlock were instrumental during "F_{ST}-gate" of 2014; I am grateful for their patience during this time. I am also thankful for the assistance I received from Wendy Easton, in obtaining permits and funding for geolocator work and both Nolan Kane and Sariel Hübner, who were instrumental at the beginning of my bioinformatics work.

All of the members of the Irwin lab have had a positive impact on my time at UBC, and include Miguel Alcaide, Alan Brelsford, Christine Grossen, Haley Kenyon, Julie Lee-Yaw, Alison Porter, Birgit Schwarz, Sampath Seneviratne and Dave Toews. I am especially thankful to Miguel and Christine who were a wealth of genomics advice. Haley and Julie became my very close friends; I am indebted to them for keeping me sane through the years. Beyond the Irwin lab I made great connections with additional members of the Biodiversity Center and am lucky to count them amongst my friends and colleagues. These people include but are not limited to Sara Miller, Brook Moyers, Richard Schuster, Diana Rennison, Kieran Samuk, Greg Owens and Thor Veen. I am especially grateful to Sara for the hours we spent commiserating together in the office and Diana, who never failed to be patient during my many trials and tribulations.

I was blessed with a great set of field and lab assistants during my PhD, including Stephanie Cavaghan, Leslie Poulson, Matt Osmond, Megan Cooling, Lauren Egli and Thomas Pierce. Each one of these people stuck with me through the ups and downs of "Team Delmore". Special thanks goes to Steph, who was the Sam to my Frodo for several years. Haley and I spent a good amount of time collecting data for Chapter 5 in the Beaty Biodiversity Museum and Field Museum of Chicago. I am thankful to the following people for showing us the ropes: Ildiko Szabo, Chris Stinson, John Bates, Ben Marks and Ben Winger.

My father was one of the most important people in my life, my biggest supporter. His loss at the beginning of my PhD threatened to upend me. I would never have completed this work without the support of my mother, Betty and brother, Tyler. Words cannot express how thankful I am to the two of you for banding together, supporting and grounding me. I never fail to be impressed by your compassion and strength. The same goes for my partner Ryan. You have stood by me for almost a decade and I cannot thank you more for your unconditional love (and devilishly good looks); whenever I faltered you were there to steady me.

Dedicated to my Father

The Road goes ever on and on

Down from the door where it began.

Now far ahead the Road has gone,

And I must follow, if I can,

Pursuing it with eager feet,

Until it joins some larger way,

Where many paths and errands meet.

And whither then? I cannot say.

– *J.R.R. Tolkien,* The Hobbit.

Chapter 1: General introduction

Seasonal migration is one of the most amazing biological phenomena in the animal world, spanning thousands of kilometers and involving billions of individuals. Humpback whales (*Megaptera novaeangliae*) undertake one of the longest journeys, migrating 25,000 km between the polar waters of the Antarctic and the tropical waters of Central America. The migration of monarch butterflies (*Danaus plexippus*) is one of the most numerous, with over one hundred million individuals moving across North America each year. And yet, many of the most iconic migrations by far are carried out by songbirds (order Passeriformes). Most migratory members of this group breed in the northern hemisphere and migrate south for the winter. They can weigh as little as 3 grams and travel up to 15,000 km. Many songbirds migrate alone and at night and return to the same breeding site each year. Common questions on this topic include, why do they migrate, what do they need to complete this journey, and how do they find their way?

My research interests are grounded in understanding the origins and maintenance of biodiversity. Accordingly, for me the most interesting question is whether seasonal migration could be responsible for generating much of the amazing diversity we see among songbirds; there are over 6,000 species in this group and more than 40% of them migrate. There are many ways that differences in migration could generate diversity. In my dissertation, I focus on whether differences in the routes taken on migration could lead to speciation. I begin this Introduction by laying the foundations of research into speciation and their connection to migration. I conclude by summarizing the main data chapters of my dissertation.

1.1 Foundations of speciation research

Formal treatment of the speciation process began during the Modern Synthesis, when population genetics was combined with Darwin's concept of natural selection and adaptation. One of the main outcomes of the Modern Synthesis was the definition of species as groups of interbreeding populations that are reproductively isolated from other such groups (biological species concept, Mayr 1942). During this time, an emphasis was also placed on studying barriers to gene flow, or *reproductive isolating barriers*, between such groups (Dobzhansky 1937a,b).

Reproductive isolating barriers can act before or after mating (pre- vs. postzygotic). Postzygotic barriers reduce the fitness of hybrids (offspring of mixed ancestry) and can be *intrinsic*, resulting from inherent incompatibilities between the genomes of pure (parental) forms or *extrinsic*, resulting from a mismatch between the environment and intermediate phenotype of hybrids (Coyne and Orr 2004; Price 2008). Initial work on isolating barriers focused on those that were intrinsic. Recently, there has been a shift to studying extrinsic barriers. This shift is likely related to an increased recognition of ecology's role in speciation (Schluter 2001; Rundle and Nosil 2005; Nosil 2012) and the observation that extrinsic barriers often act early in divergence. Accordingly, they may be more important to the process of speciation (Schluter 2009; Schluter and Conte 2009; Wolf et al. 2010; Seehausen et al. 2014).

Around the same time that ecology's role in speciation began receiving more attention in the literature, there was a re-evaluation of the biological species concept. The original formulation of this concept emphasized the independent nature of species, with species viewed as co-adapted gene complexes resistant to the influence of gene flow. Wu challenged this notion (Wu 2001; Wu and Ting 2004), arguing that only a subset of genes may be responsible for reproductive isolation. Accordingly, gene flow throughout the rest of the genome may occur

without degrading species (i.e., species need not be viewed as independent units). Many authors supported Wu's argument, including Coyne and Orr who summarized work on speciation in 2004 and modified the biological species concept to permit gene flow, requiring "substantial but not complete reproductive isolation" between groups. This adjustment stimulated a proliferation of research on the *genetic basis of speciation*. Work on this topic has been aided by the advent of next generation sequencing techniques. These techniques have parallelized DNA sequencing, allowing the production of millions of reads (sequences) in a short amount of time (Metzker 2010; Stapley et al. 2010b; Davey et al. 2011).

Research into reproductive isolating barriers and the genetic basis of reproductive isolation has been facilitated in large part by *hybrid zones*, the areas where distinct groups meet and interbreed. Barton and Hewitt were among the first authors to recognize the power of hybrid zones, developing equations to estimate the strength of selection acting against hybrids and the number of loci involved in reproductive isolation (Barton and Hewitt 1985; Hewitt 1988; Barton and Gale 1993). Their work on the tension zone model of hybrid zone stability also highlighted the impact that selection at a single locus could have on neighbouring loci, reducing introgression over large regions of the genome (Barton and Hewitt 1985). More recently, authors have made use of variable genotypes and phenotypes in hybrid zones to measure the relationship between phenotypes, genotypes and fitness and conduct genetic mapping (Rieseberg and Buerkle 2002; Buerkle and Lexer 2008; Gompert and Buerkle 2011). It should also be noted that speciation concerns the evolution of reproductive isolation. Accordingly, research on this topic should focus on taxa that are in the process of becoming reproductively isolated (Via 2009), making hybrid zones, where hybridization continues to occur, the perfect arena for work on speciation.

1.2 Focus on birds and seasonal migration

Many of the most influential authors in speciation were inspired by birds; Darwin's ideas on natural selection originated partially from artificial selection experiments conducted on pigeons and Mayr's ideas on reproductive isolation derived in large part from his observations of avian geographic distributions. Despite this attention, we still have large gaps in our knowledge of avian speciation. For instance, there is good evidence that differences in resource use serve as isolating barriers between Darwin's finches (Grant and Grant 1997) but we lack this resolution in most other systems. A focus on extrinsic postzygotic barriers might be exactly what we need, as both prezygotic and intrinsic postzygotic barriers have been shown to play a limited role in birds (Grant and Grant 1992; Price and Bouvier 2002; Rabosky and Matute 2013). Research on the genetic basis of reproductive isolation in birds is also in its infancy. For instance, the only reference genomes available prior to 2010 for birds were the chicken and zebra finch. Both groups are model organisms with limited relevance to speciation (e.g., the chicken does not sing and neither migrate [see relevance below]).

There has been a great deal of recent interest in the role seasonal migration could play as an extrinsic barrier to gene flow between birds. This suggestion stems from the observation that many taxa from this group form migratory divides on their breeding grounds. Migratory divides are contact zones between divergent populations that breed adjacent to one another but use different migratory routes (Helbig 1991; Dallimer et al. 2003; Bearhop et al. 2005; Irwin and Irwin 2005; Veen et al. 2007; Bensch et al. 2009; Møller et al. 2011). Migratory routes are largely genetically determined in many groups, including songbirds (Pulido 2007; Liedvogel et al. 2011) and often involve navigation around unsuitable habitat (e.g., mountain chains and

deserts that would be difficult to navigate over and provide limited habitat to stop and forage on during migration). Accordingly, it has been predicted that (1) hybrids will use intermediate migratory routes and (2) that these routes will be inferior to those of parental forms (e.g., cause them to migrate over rather than around unsuitable habitat). Under this scenario, hybrids will be selected against and differences in migratory orientation will reduce gene flow between groups. Note that the former predictions assume that migratory orientation is inherited additively, which has been shown in orientation cage experiments (Helbig 1996).

Empirical tests of the migratory divide hypothesis have been limited primarily to birds in the western Palearctic and often to laboratory studies (e.g., Helbig 1991; Veen et al. 2007; Bensch et al. 2009). Given the large number of taxa that form migratory divides (Irwin and Irwin 2005; Møller et al. 2011; Rohwer and Irwin 2011) and potential importance of extrinsic postzygotic barriers to avian speciation, differences in migratory orientation could serve as powerful reproductive isolating barriers.

1.3 Study system: the Swainson's thrush

The first three data chapters of my dissertation focus on the Swainson's thrush (*Catharus ustulatus*). The Swainson's thrush is a migratory songbird that breeds in the boreal forests of North America and riparian woodlands along the Pacific Coast. This species includes four morphological subspecies that occur in two mitochondrial clades: inland, olive-backed and coastal, russet-backed thrushes (Ruegg and Smith 2002). Paleodistribution models and phylogenetic analyses suggest that these groups were isolated in different refugia during the last glacial maximum (Ruegg et al. 2006a). They have since expanded their ranges northward and currently form a hybrid zone along the Coast Mountains of British Columbia (Ruegg 2008).

Substantial reproductive isolation has been documented between inland and coastal thrushes; clines in morphology, plumage and genetic traits are narrow and coincident across the hybrid zone and population densities at its center are low (Ruegg 2008). Differences between coastal and inland Swainson's thrushes have been described in their songs (Ruegg et al. 2006b) and the abiotic elements of their breeding environments (Ruegg et al. 2006a); selection associated with these differences could be contributing to reproductive isolation between the groups. Differences in migratory routes have also been suggested: data from band recoveries and mitochondrial haplotypes sampled at stopover and wintering sites suggested that the inland form migrates along a southeastern route to its wintering grounds in South America while the coastal form migrates along a southern route to its wintering grounds in southern Mexico and Central America (Figures 1 and 2 in Ruegg and Smith 2002). These differences make the Swainson's thrush an ideal system to study the migratory behaviour of hybrids and migration's possible role in reproductive isolation. If hybrids take intermediate routes across the deserts and mountains of the American West, they may be less successful than the pure coastal and inland forms, reducing gene flow between the groups.

1.4 Research conducted for this dissertation

A variety of methods have been used to identify migratory divides, including band recovery data (Chamberlain et al. 2000; Ruegg and Smith 2002), orientation experiments (Helbig 1991; Dallimer and Jones 2002), genetic (Ruegg and Smith 2002; Davis et al. 2006) and isotopic markers (Veen et al. 2007; Bensch et al. 2009). The latter two methods are often considered the most powerful, as information on wintering, breeding and/or stopover sites can be obtained by capturing a bird only once (Webster 2002). Nevertheless, these methods are indirect (e.g., the

wintering location of a breeding bird must be inferred based on the isotopic signatures in its feathers) and often limited to providing broad scale inferences (e.g., eastern vs. western North America, Lovette et al. 2004).

Light-level geolocators could represent a solution these problems. These devices record light intensity at specific time intervals and were recently miniaturized, allowing their use on songbirds. They are fitted to birds on the breeding grounds and retrieved the following year, when light intensity data can be used to infer longitude and latitude. One disadvantage of this technology is that, similar to band recovery data, it requires the recapture of birds. Nevertheless, far more detailed data are obtained with this technology (e.g., only two location fixes are obtained with band recovery data). I harnessed the power of geolocators in Chapter 2, attaching devices to birds from pure populations adjacent to the hybrid zone between inland and coastal thrushes (Figure 1.1). The main objective of this chapter was to confirm the existence of a migratory divide between these groups using tracking data. Recall that this divide was originally identified using band recovery data and genetic markers (Ruegg and Smith 2002). In addition, very little of the data from this previous analysis came from directly adjacent to the hybrid zone, where differences in migration would have the greatest influence on reproductive isolation.

Far less work has been done to actually evaluate the migratory divide hypothesis (i.e., test the predictions that hybrids will take intermediate and inferior routes on migration). The best evaluations of this hypothesis come from work conducted on European blackcaps (*Sylvia atricapilla*) using orientation cages. For example, two populations of blackcap form a divide in central Europe; one migrates southwest to Portugal and Spain and the other southeast to eastern Africa. Helbig (1991) bred F_1 hybrids from parental individuals and assayed their orientation using Emlen funnels. In accordance with the migratory divide hypothesis, hybrids oriented in

directions intermediate to parental forms. Helbig (1991) hypothesized that this orientation would be inferior as it would bring hybrids over the Alps, the Mediterranean Sea and the Sahara Desert.

As alluded to in the introduction to Chapter 2, orientation cages provide only limited information on migration; researchers are hoping to quantify migratory restlessness (or *Zugunruhe*) in these cages. In reality, there is a lot of variation in the movement behaviour that birds exhibit in cages and much of it is likely unrelated to migration. In addition, these experiments are often run for only one to two weeks in the fall. To evaluate the migratory divide hypothesis directly, free-flying birds must be tracked over the entire annual cycle. Accordingly, in Chapter 3, I attached geolocators to birds at the center of the hybrid zone between inland and coastal Swainson's thrushes (Figure 1.1). The main objective of this chapter was to evaluate the key prediction from the migratory divide hypothesis, that hybrids will take intermediate routes, using tracking data. I also used (1) cline theory to obtain an estimate of selection against hybrids and (2) inferences from the routes taken by hybrids to evaluate whether these routes were inferior to those taken by parental forms.

In Chapter 4 I switched from this largely phenotypic work to begin examining the genomic basis of reproductive isolation between inland and coastal Swainson's thrushes. Following Wu's work in the early 2000s, a series of genome scans were conducted to quantify differentiation between divergent taxa. Most scans documented highly heterogeneous landscapes of differentiation (e.g., Ellegren et al. 2012; Nadeau et al. 2013; Renaut et al. 2013) and there is currently a push to identify the evolutionary processes generating these patterns. Two main models are at the forefront of this research. The first model derives from Wu's work, suggesting that areas of increased differentiation include loci involved in reproductive isolation. Any area that is not diverged beyond the background level must have been homogenized by gene flow

(divergence-with-gene-flow model; Nosil et al. 2009; Feder et al. 2012; Via 2012). A second model was proposed more recently, and suggests that gene flow need not be involved. Instead, variable selection across the genome (directional or background) could generate these patterns (selection-in-allopatry model; Noor and Bennett 2009; Cruickshank and Hahn 2014).

Work for Chapter 4 began by assembling a draft reference genome for the Swainson's thrush. This gave me more resolution for subsequent analyses, in which I obtained re-sequencing data from populations adjacent to the hybrid zone between inland and coastal thrushes (Figure 1.1). One of the main objectives of this chapter was to use patterns of genetic differentiation and within population variation to evaluate the two main models of genomic heterogeneity introduced above (divergence-with-gene flow versus selection-in-allopatry). The second objective of this chapter was to connect this work to the previous two chapters by generating lists of candidate genes associated with migration and determining whether these genes showed evidence of being under selection. If differences in migration are contributing to reproductive isolation between thrushes, these areas may show signs of selection.

Finally, in Chapter 5 I extend my work on the Swainson's thrush to evaluate the migratory divide hypothesis on a broader scale. A similar, although more qualitative, approach has been taken by other authors (Irwin and Irwin 2005; Møller et al. 2011; Rohwer and Irwin 2011). Irwin and Irwin (2005), for instance, reviewed the literature to identify birds that breed in Siberia and winter in southern Asia. These taxa encounter the Gobi Desert and Tibetan Plateau on migration. Irwin and Irwin (2005) identified 97 long-distance migrants, 82 of which use solely western or eastern routes around the high-elevation deserts of the Tibetan Plateau. Seven of the remaining species had different subspecies going west and east around the Plateau, suggesting a role for migration in reproductive isolation. Together, these studies established the

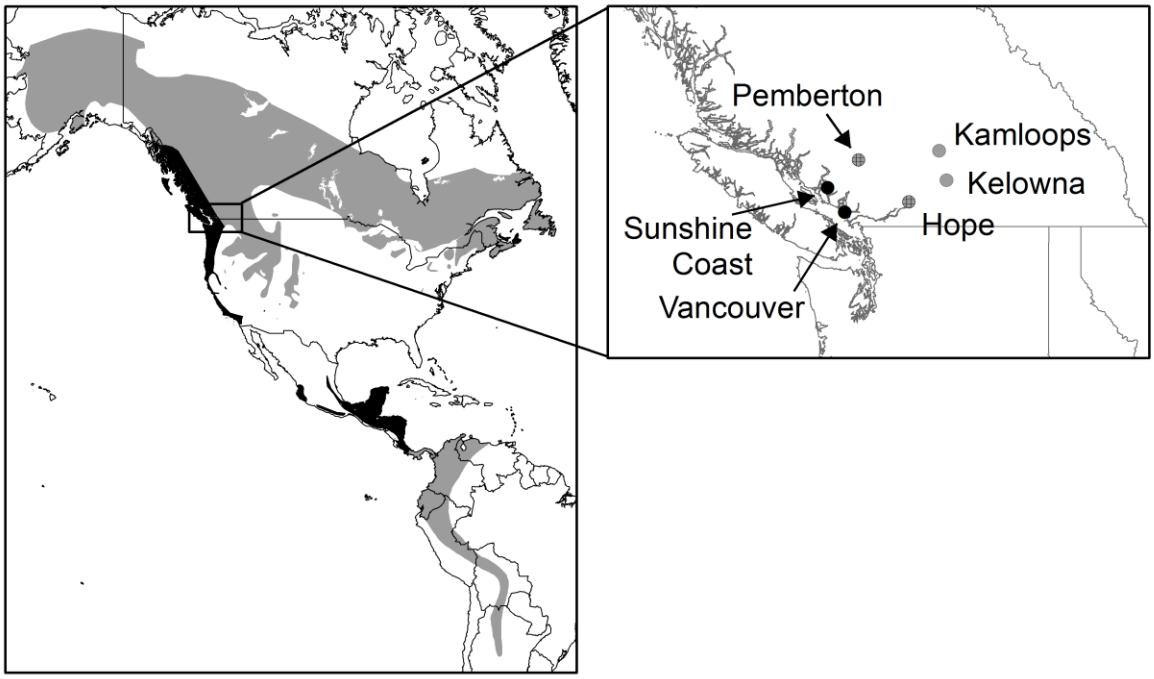
plausibility that differences in orientation could serve as major isolating barriers between birds. Unfortunately none of the previous studies included a rigorous statistical analysis or a quantitative measure of reproductive isolation. The main objective of this last chapter was to use a comparative and quantitative approach to evaluate the migratory divide hypothesis on a broader scale. I used divergence in phenotypic traits involved in species recognition as a proxy for reproductive isolation (Martin et al. 2010; Tobias et al. 2010; Weir and Wheatcroft 2011) and compared divergence between closely related species of birds from North America that do or do not form divides. If differences in migration contribute to speciation, I predicted that divergence would be greater among species pairs that form divides.

1.5 Summary

In summary, my dissertation uses modern tracking and genetic techniques to explore the following questions related to the role of migration as a reproductive isolating barrier and the genetic basis of reproductive isolation.

1. Does a migratory divide exist between Swainson's thrushes?
2. Do hybrids take intermediate and inferior routes in the migratory divide between Swainson's thrushes?
3. What is the genetic basis of reproductive isolation between Swainson's thrushes?
4. Among migratory birds in North America, are migratory divides associated with greater phenotypic divergence?

Figure 1.1 Geographic range of inland and coastal Swainson's thrushes and population samples included in this dissertation. The coastal subspecies group's range is shown in black and the inland group's range in gray. Samples from the Sunshine Coast and Vancouver (black circles) and Kamloops and Kelowna (gray circles) occur adjacent to the hybrid zone. Pemberton and Hope (gray circles with hatched black lines) occur at the center of the hybrid zone between Swainson's thrushes.



Chapter 2: Dramatic intraspecific differences in migratory routes, stopover and wintering sites revealed using light-level geolocators[1]

2.1 Introduction

Migratory divides are contact zones between divergent populations that breed adjacent to one another but use different routes to reach their wintering grounds (Helbig 1996; Bensch et al. 1999; Irwin and Irwin 2005). It is generally agreed that these divides formed following secondary contact between populations that were isolated in different glacial refugia during the Pliocene and/or Pleistocene (Berthold 2001; Irwin and Irwin 2005; Møller et al. 2011; Rohwer and Irwin 2011). During this time, these populations diverged in multiple traits, including migratory orientation. The migratory routes these populations use today likely reflect the colonization routes their ancestral forms employed following the last glacial maximum (Rappole 1995; Berthold 2001; Newton 2008).

Migratory divides have been described in many taxonomic groups (Gagnaire et al. 2009; McDevitt et al. 2009; Altizer and Davis 2010), although the majority of research has focused on songbirds. Of particular interest is the role differences in migratory behaviour play in maintaining genetic differentiation, local adaptation and reproductive isolation (Bearhop et al. 2005; Boulet et al. 2006; Bensch et al. 2009; Rolshausen et al. 2009). Research has also focused on the relevance of migratory divides to our understanding of the ecology and conservation of migrants; populations in divides encounter different sets of ecological conditions during

[1] A version of this chapter has been published: Delmore KE, Fox JW, Irwin DE. 2012. Dramatic intraspecific differences in migratory routes, stopover and wintering sites revealed using light-level geolocators. *Proceedings of the Royal Society B: Biological Sciences* 279: 4582-4589.

migration and on the wintering grounds (Kelly and Hutto 2005). These differences could affect their reproduction on the breeding grounds (Marra et al. 1998; Norris et al. 2004) and suggest that these populations should be considered independent management units when establishing conservation strategies (Boulet et al. 2006; Faaborg et al. 2010b).

To date, migratory divides have been described using band recovery data and/or biological markers (Boulet et al. 2006; Bensch et al. 2009; Irwin et al. 2011). Band recovery data are often limited by sample size (Webster et al. 2002), can fail to include individuals from the area of interest (e.g., directly adjacent to a migratory divide) and provide only two location fixes over the entire annual cycle (initial capture and recapture). Biological markers are often restricted to describing broad scale patterns and are indirect; individuals are not followed over the entire annual cycle (Webster et al. 2002; Lovette et al. 2004; Hobson 2005; Smith et al. 2005). In contrast, light-level geolocators, which are archival tags that record light intensity at specific time intervals, provide a means to directly track individuals. These devices are attached to birds on the breeding grounds and retrieved the following year. Through astronomical algorithms as well as atmospheric and movement models, the archived light intensity data can be analysed to produce positional data (Hill 1994). Geolocators were recently miniaturized, permitting their use on songbirds (Stutchbury et al. 2009) and providing the opportunity to describe migratory divides using a more direct method.

The Swainson's thrush (*Catharus ustulatus*) is a Neotropical migrant that breeds throughout the boreal forests of North America and winters in southern Mexico, Central and South America (Mack et al. 2000). Two subspecies groups have been described: the russet-backed, coastal group and olive-backed, inland group (Ruegg and Smith 2002). Ruegg (2008) described a hybrid zone between these groups along the Coast Mountains of British Columbia.

Data from band recoveries and mitochondrial haplotypes sampled at stopover and wintering sites suggested that these groups form a migratory divide (Ruegg and Smith 2002). The band recovery data reported by Ruegg and Smith (2002) contained very little data from British Columbia, where the hybrid zone occurs (two birds from southeast BC, far east of the hybrid zone and one from coastal BC, west of the hybrid zone, see Figure 2 in Ruegg and Smith 2002). In addition, no band recovery data were included from the interior western United States. This region forms a large part of the inland groups range and raises the possibility that breeding populations just east of the hybrid zone migrate through the interior western United States, a route closely parallel to that of the coastal group. If this were the case, there would not be a strong migratory divide between birds on either side of the Swainson's thrush hybrid zone. We tested whether there is actually a narrow migratory divide here, using light-level geolocators attached to birds at the western and eastern edges of the hybrid zone between inland and coastal Swainson's thrushes.

2.2 Methods

We attached 39 geolocators to male Swainson's thrushes at three sites in British Columbia, Canada in June 2010 (Sunshine Coast [coastal], $n = 10$; Vancouver [coastal], $n = 9$; Kamloops [inland], $n = 20$, Figure 1.1). Birds were caught using song playback and mist nets and fitted with a Canadian Wildlife Service aluminum band. Body mass and wing, tarsus and tail length of each bird were measured. We also photographed each bird and obtained a tail feather (R4, the fourth right rectrix) for later genetic analyses. We used Mk12S geolocators (0.9 g, stalk length = 15 mm; British Antarctic Survey [BAS], Cambridge, United Kingdom) and attached these devices to birds using Rappole-Tipton leg-loop backpack harnesses (Rappole and Tipton 1991) made of Teflon ribbon (Bally Ribbon), EPDM or silicone cord (Budlar; 0.2 – 0.3 g). Total attachment

weight was less than 4% the average body mass of Swainson's thrushes (30.5 ± 1.5 g, Delmore unpub. data).

We used BASTrak software (BAS) to download light intensity data and estimate daily latitude and longitude at local noon. Mk12S geolocators measure light intensity every minute and record the maximum measurement every two minutes. Light intensity is recorded on an arbitrary scale between 0 and 64 and is used in combination with a Sun elevation (altitude) angle to define light transitions (i.e., sunrise and sunset). Lower values of light intensity are recommended for most birds, especially those spending time in shaded environments. Accordingly, we used a value of 1 and determined the equivalent Sun elevation calibration value from data when birds were on the breeding grounds (i.e., a known location). Calculated Sun elevation angles ranged between -4.5 and -3 degrees and we used the average of these values (-3.7) for the analysis of the entire dataset (see Appendix A.1 for one exception). We rejected transitions with high levels of noise and evidence of shading (34% of days). We also rejected latitudinal estimates within 15 days of the fall and spring equinoxes as well as obvious outliers, most of which occurred near the equinoxes (3% of days). Latitude and longitude can be estimated at local midnight and local noon. Assuming the birds only migrated at night, we used noon location fixes only (without movement compensation) as these should be more accurate, being uncomplicated by analysis errors induced by movement (e.g., movement in longitude causes an apparent increase or decrease in day length). We evaluated the accuracy of our estimates using the standard deviation of longitude and latitude while birds were on the breeding grounds and the distance between the true deployment location and the location estimated using geolocators.

We plotted daily estimates of latitude and longitude in ArcGIS (Esri) and calculated average locations for days when we assumed birds were stationary (at stopover sites on the

breeding or wintering grounds). Following Stutchbury et al. (2009), we defined stopover sites as locations where birds remained stationary for ≥ 2 days. We mapped points that fell over water to the closest point on land and used the most direct route when connecting points. It should be noted that by using the most direct route, we are likely underestimating the distance travelled by each bird. In addition, we were unable to obtain latitude estimates during a large portion of the fall migration due to the equinox (see Results). Rather than connecting points before and after the equinox using straight lines, we used information from estimates of longitude to draw our lines. More specifically, we identified three stopover sites employed by birds using longitude data during this period and used this information to interpolate between points.

2.3 Results

We recovered and successfully downloaded data from 10/39 devices in 2011. The light stalk fell off one of these devices on July 27, 2010 and was not used in our analyses. The battery on one of the devices stopped working on May 16, 2011. Fortunately the bird had almost completed its spring migration, allowing us to include this device in our analyses. Average standard deviation for location estimates on the breeding grounds was 106.77 km for latitude and 37.67 km for longitude. Average distance between the true deployment location and the estimate from geolocator data was 117 km.

Figure 2.1 depicts the annual cycle of the thrushes we recaptured. Coastal birds migrated along the western coast of North America on fall and spring migration and wintered in southern Mexico and Central America (Guatemala and Honduras). Inland birds employed more eastern routes, passing over the Rocky Mountains and through the central United States. These birds migrated over the Gulf of Mexico on fall migration, wintered in South America (Colombia and

Venezuela), and migrated around the Gulf through Central America and Mexico on spring migration (Appendices A.2 – A.4 include graphs for individual birds and daily estimates of longitude and latitude ± standard deviation).

All of the coastal birds employed a common stopover site in the Sierre Madre mountain ranges of southern Mexico on spring migration. These birds spent a prolonged period of time at this stopover site (average = 19 days, range = 8 – 29 days, Figure 2.1b). Inland birds also employed common stopover sites for extended periods of time. On fall migration, these birds stopped at a site north of the Gulf of Mexico, in Alabama (average = 20 days, range = 11 – 29 days, Figure 2.1a). On spring migration, these birds stopped at sites south of the Gulf of Mexico, in Panama (average = 11 days, range = 7 – 16 days) and Guatemala (average = 11 days, range = 3 – 17 days), and at one site north of the Gulf, in Texas (range = 9 – 15 days, average = 13 days, Figure 2.1b).

Latitude is estimated using day length. There is little variation in this variable around the equinoxes and, as a result, we were unable to estimate latitude for ~ 30 days around each of these periods. These missing days are illustrated in Figure 2.1 as dashed lines and overlap substantially with fall migration routes. One consequence of this overlap is that we are missing data from two of the inland birds at the first stopover site described above, on the northern coast of the Gulf of Mexico in Alabama. Estimates of longitude are unaffected by the equinoxes, however, and a plot of longitude by date suggests that these birds may have spent time at the same stopover site, as all four birds remained at roughly the same longitude in late October (~ 85º, Figure 2.2). This plot also suggests that all four coastal birds spent a substantial amount of time at one or two stopover sites between 106º and 96º longitude on fall migration. These sites may be located in

the Mexican Monsoon region of western North America, where many western breeding birds stop between the months of July and October to molt (Rohwer et al. 2005).

Table 2.1 summarizes differences in the timing and distance travelled by birds from coastal and inland subspecies groups. Spring migration was shorter in duration than fall migration for birds from both groups (paired t-test, t = -3.90, df = 12.7, P = 0.0016). Significant differences between inland and coastal birds were observed in the distance travelled on both spring and fall migration, with inland birds migrating longer distances on each leg. Significant differences were not observed between these groups in any of the other variables measured, including departure and arrival date, duration of migration and number of days spent at stopover sites. Nevertheless, the data suggest that coastal birds left their breeding grounds earlier than inland birds. Coastal birds also appear to have spent less time on spring migration, leaving their wintering grounds later than inland birds but arriving on the breeding grounds earlier.

The tracks of individual birds revealed additional details from the annual cycle of these groups. First, three of the coastal birds employed two distinct sites on the wintering grounds, moving from their first site at the end of December. Two birds moved east from Honduras to Guatemala and southern Mexico (Figure 2.3ab; Appendix A.3ab), one bird moved west from southern Mexico to Guatemala (electronic supplementary material Appendix A.3gh). Average distance between these sites was 421 km. Second, as mentioned briefly above, two of the inland birds appear to have flown over the Gulf of Mexico on fall migration, leaving from Alabama and arriving in Honduras. These birds flew around the Gulf on spring migration, taking a land route through Central America and Mexico (Figure 2.3cd; Appendix 4cd). It is likely that the other two inland birds followed a similar path, as these birds appear to have spent time at the same stopover site in the southeastern United States (described above, Figure 2.2).

2.4 Discussion

Data from light-level geolocators deployed in our study revealed dramatic differences in the migratory routes, stopover sites and wintering grounds employed by inland and coastal Swainson's thrushes. To the best of our knowledge, this is the first study in which data from individual birds over their entire annual cycle have been used to characterize long distance migration routes across a migratory divide. Our results confirm and expand on conclusions presented in Ruegg and Smith (2002) based on band recovery data and genetic markers. More specifically, geolocators provided location estimates for most days over the entire annual cycle of each bird. This additional resolution allowed us to document several novel aspects of the routes employed by Swainson's thrushes, including the circum-Gulf route employed by inland birds and long-term stopover sites employed by both groups. Geolocator data also suggest that that the trans-Gulf route employed by these birds is not as far east as was suggested with the banding data, which only provides two locations per bird; geolocators show birds crossing between Alabama and the Yucatan, whereas the banding data suggested crossing between Florida and Columbia. Finally, the banding data presented by Ruegg and Smith (2002) included long distance movements from the breeding range to Central America for only two birds. The present study provides fives case, and tremendous detail on those movements. Below, we discuss the relevance of our results to understanding the evolution, ecology and conservation of migratory species.

2.4.1 Evolution of migratory routes and their role as reproductive isolating barriers

Migratory routes are believed to reflect historical range shifts, with populations migrating along the same routes their ancestral forms used to expand out of glacial refugia (Rappole 1995; Berthold 2001; Irwin and Irwin 2005; Newton 2008). Our results partially support this suggestion in Swainson's thrushes. Population genetic analyses and paleodistribution modelling suggested that inland and coastal thrushes were isolated in separate eastern (inland) and western (coastal) refugia during the last glacial maximum. Both groups expanded northward out of these refugia; once the inland group reached the eastern boreal forest it expanded west into the western boreal forest (Ruegg et al. 2006a). Coastal thrushes migrated along a north – south axis in our study, supporting the suggestion that they are retracing their post-glacial colonization routes. Routes employed by inland thrushes on fall migration also support this suggestion; these birds migrated southeast to their ancestral range in the eastern United States before continuing south to their wintering grounds.

Migratory routes have also been implicated in the maintenance of reproductive isolation between populations. Specifically, migratory orientation is largely genetically determined in many groups, including songbirds (Pulido 2007) and often involves navigation around geographic areas that are difficult to migrate over and/or have little suitable habitat for refuelling (Alerstam 2001). Accordingly, hybrids between populations at these divides are expected to employ intermediate routes that will be inferior to those of parental forms (Helbig 1996; Irwin and Irwin 2005; Bensch et al. 2009; Rohwer and Irwin 2011). Our results support this suggestion in Swainson's thrushes. Inland and coastal thrushes tracked in our study employed divergent migratory routes that took them east or west of several large geographic features in western North America, including three mountain ranges (Cascade, Sierra Nevada and Rocky Mountains)

and deserts in the southwestern United States. Inland and coastal thrushes are genetically differentiated from one another and an analysis of hybrid populations suggests that they exhibit some degree of reproductive isolation (Ruegg and Smith 2002; Ruegg 2008).

2.4.2 Contributions to our knowledge of the non-breeding ecology of migratory species

Migratory connectivity has been defined as the link between an individual's breeding and wintering grounds (Webster et al. 2002). This link appears to be relatively strong for inland thrushes; these birds migrated to sites within 2 degrees longitude in Columbia and Venezuela. Connectivity was lower in coastal birds; these birds migrated to sites within 8 degrees longitude in southern Mexico, Honduras and Guatemala. Three coastal birds employed two wintering sites, contributing to lower levels of connectivity in this group. This behaviour likely represents an example of intratropical migration (Faaborg et al. 2010a; Heckscher et al. 2011). To the best of our knowledge, this is only the second time this behaviour has been documented in a Neotropical migrant; Hecksher et al. (2011) documented a similar pattern in Veeries, in which birds wintering in in the Amazon Basin moved to a second site between the months of January and March. Together, these results suggest that migratory connectivity may not be as straightforward as was once believed; studies of this phenomenon should not be constructed to assign birds to single wintering sites.

Two additional details relevant to our understanding of the non-breeding season of Swainson's thrushes were revealed in our study. First, at least two of the inland birds we tracked exhibited seasonal differences in the migratory routes they employed, using a trans-Gulf route on fall migration and a more western circum-Gulf route on spring migration. This behaviour has been termed loop migration and has been described using data from bird monitoring stations and

band recoveries for other Neotropical migrants (*Selaphorus* hummingbirds, Phillips 1975; blackpoll warblers, Hunt et al. 1999). This behavior was also described in two recent studies employing light-level geolocators in the western Palearctic (European hoopoe, Gauthreaux 1999; red-backed shrike, Bächler et al. 2010). Very little is known about loop migration but it is likely related to variation in food availability and/or prevailing winds (Phillips 1975; Hunt et al. 1999; Faaborg et al. 2010a). Second, individuals from from both inland and coastal groups employed similar long-term stopover sites. These sites were located near large geographic barriers, including the Gulf of Mexico and the Sierra Madre mountain ranges and were likely used to acquire the resources necessary to cross these barriers; birds spent prolonged periods of time at these sites and crossed these barriers immediately after. Previous studies employing data from bird monitoring stations and radar technology have highlighted the importance of these sites for migratory songbirds (Felger and Wilson 1994; Gauthreaux 1999; Moore 1999; Skagen et al. 2005; Gauthreaux et al. 2006). Veeries and Wood thrushes fit with geolocators in eastern North America also relied on similar stopover sites around the Gulf of Mexico (Stutchbury et al. 2009; Heckscher et al. 2011).

2.4.3 Relevance for establishing conservation strategies for migratory species

Details from the non-breeding season of Swainson's thrushes described above can be used to inform year-round conservation strategies for songbirds. For example, the long-term stopover sites identified are likely important for completing migration along eastern and western routes; all of the birds from each group stopped at these sites and in many cases spent more than a week at each. We can use these data to establish sets of stopover sites along these routes that should be the focus of conservation efforts. Data collected in our study can also be used to inform

conservation strategies in the future. For example, migratory connectivity appears to be relatively high within subspecies groups. If a decline is observed in one group we should focus management efforts on the specific breeding, wintering and stopover sites employed by individuals from this group. It should be noted that the data we presented are only from one species. Nevertheless, similar east-west genetic differentiation has been described in several other species of North American songbirds and data from biological markers suggests that many of these populations may follow similar routes (Lovette et al. 2004; Smith et al. 2005; Boulet et al. 2006; Irwin et al. 2011). Accordingly, we should be able to apply results obtained in this single species study to the conservation of many migratory species.

Avian migration is perhaps the most geographically widespread of all biological phenomena, as it ties together the ecology of locations separated by great distances, often on different continents. Migratory divides are particularly interesting in this regard, because they show that neighbouring breeding populations can be widely separated geographically at other times throughout the year. As discussed above, these large migratory differences can have importat consequences for evolution (e.g., by contributing to reproductive isolation), ecology (e.g., because the two groups depend on different wintering environments), and conservation (e.g., by revealing important stopover sites). Geolocators now provide us with a tool to reconstruct detailed pathways of individual birds, allowing researchers to examine a wide variety of questions in these fields. We anticipate that great advances in understanding patterns of migratory connectivity will soon be made by combining analyses of trackways employed by individual birds with population-level analyses of genetic and feather isotope variation.

Table 2.1 Details on the fall and spring migration of nine Swainson's thrushes tracked from pure populations. Departure and arrival dates are shown along with total duration, distance travelled, days spent at stopover sites, migration pace and flight speed. Colours correspond to those used in Figure 2.1 (inland subspecies group: red, orange, yellow, green; coastal subspecies group: blue, purple, pink, turquoise, maroon). Averages for each subspecies group are shown in bold. Significant results from paired t-tests comparing values for these groups are also bolded.

	Fall migration							Spring migration						
	Departure	Arrival	Duration (days)	Stopover (days)	Distance (km)	Pace (km/day)	Speed (km/day)	Departure	Arrival	Duration (days)	Stopover (days)	Distance (km)	Pace (km/day)	Speed (km/day)
Coastal														
Maroon	12 Jul	17 Oct	94	83	3470	37	315	16 Apr	01 Jul	46	31	4253	92	284
Turq[a]	25 Jul	18 Oct	85	54	4366	51	141	06 Apr	28 May	52	39	4441	85	342
Blue	28 Jul	14 Oct	78	71	3759	48	537	18 Apr	23 May	36	21	4564	127	304
Pink[b]	24 Jul	20 Oct	88	72	4466	51	279	01 Apr						
Purple	26 Jul	02 Oct	68	35	4067	60	123	11 Apr	17 May	36	28	4041	112	505
Average	**23 Jul**	**14 Oct**	**83**	**63**	**4026**	**49**	**279**	**10 Apr**	**25 May**	**44**	**33**	**4325**	**104**	**359**
Inland														
Red	31 Jul	27 Nov	119	98	5468	46	260	26 Mar	09 Jun	75	57	5552	74	308
Green	20 Jul	03 Nov	96	85	6081	63	553	04 Apr	06 Jun	63	38	6333	101	253
Orange	31 Jul	17 Oct	78	66	5519	71	460	07 Apr	22 May	45	31	6083	135	435
Yellow	28 Aug	08 Oct	41	32	5122	125	569	10 Apr	21 May	41	32	5532	135	615
Average	**06 Aug**	**29 Oct**	**84**	**70**	**5547**	**76**	**461**	**04 Apr**	**30 May**	**57**	**44**	**5875**	**111**	**403**
P-value	0.073	0.18	0.95	0.66	**0.0008**	0.13	0.13	0.20	0.43	0.18	0.21	**0.0001**	0.71	0.66

[a] Turquoise
[b] battery stopped working May 16, 2011

Figure 2.1 Annual cycle of nine Swainson's thrushes tracked from pure populations. Fall (a) and spring (b) migration are shown. Routes for thrushes from the inland subspecies group are shown in warm colours (red, orange, yellow, green); routes for thrushes from coastal subspecies group are shown in cool colours (blue, purple, pink, turquoise, maroon). Dashed lines link locations where latitude could not be estimated around the equinox periods. Long-term stopover sites are shown in the United States (Alabama, diamond; Texas, pentagon), Central America (Panama, circle; El Salvador and Guatemala, triangle) and southern Mexico (square).

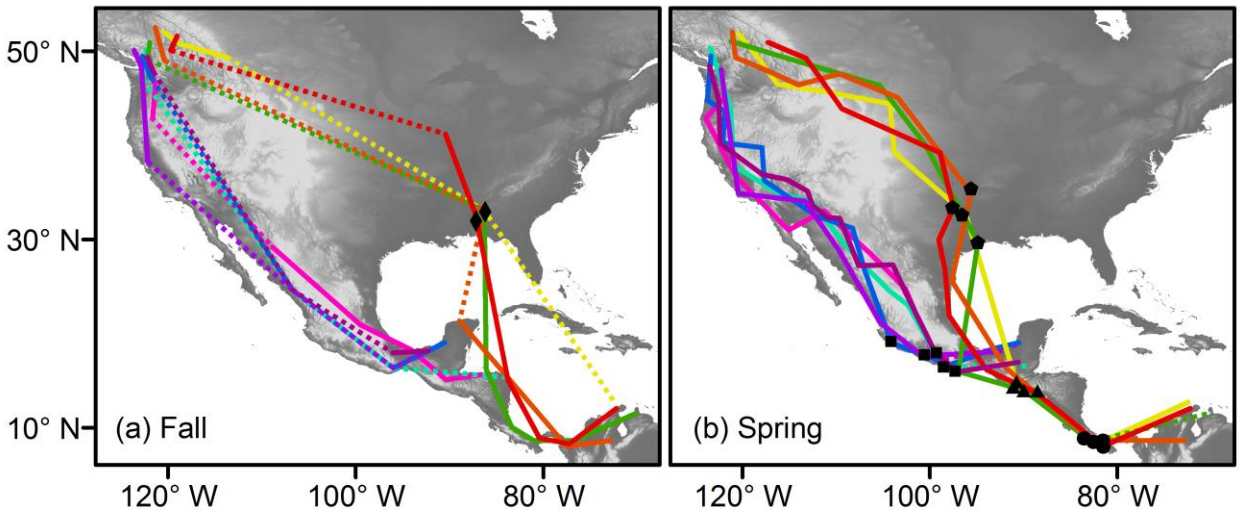


Figure 2.2 Change in longitude during the annual cycle of nine Swainson's thrushes tracked from pure populations. Colours correspond to those used in Figure 2.1, with data for thrushes from the inland subspecies group shown in warm colours (red, orange, yellow, green) and from the coastal subspecies group shown in cool colours (blue, purple, pink, turquoise, maroon).

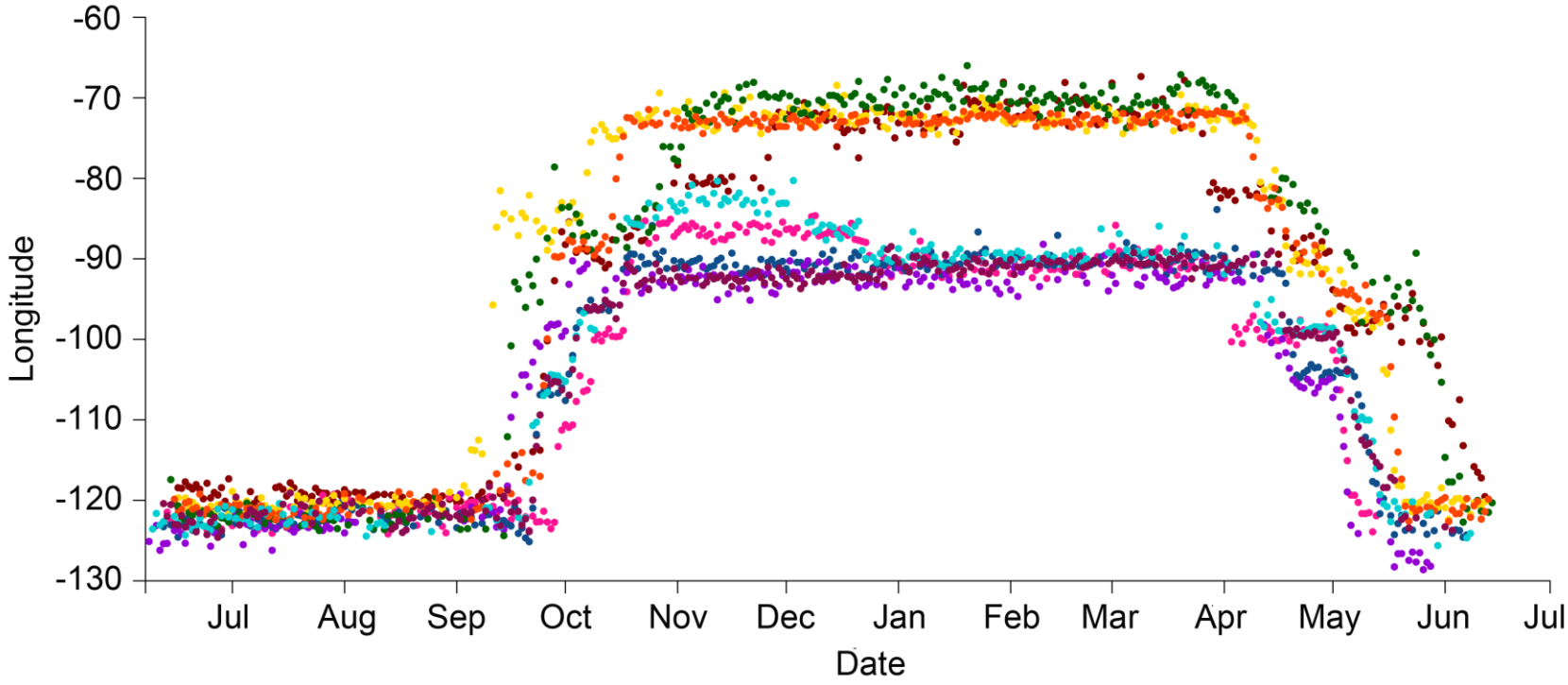
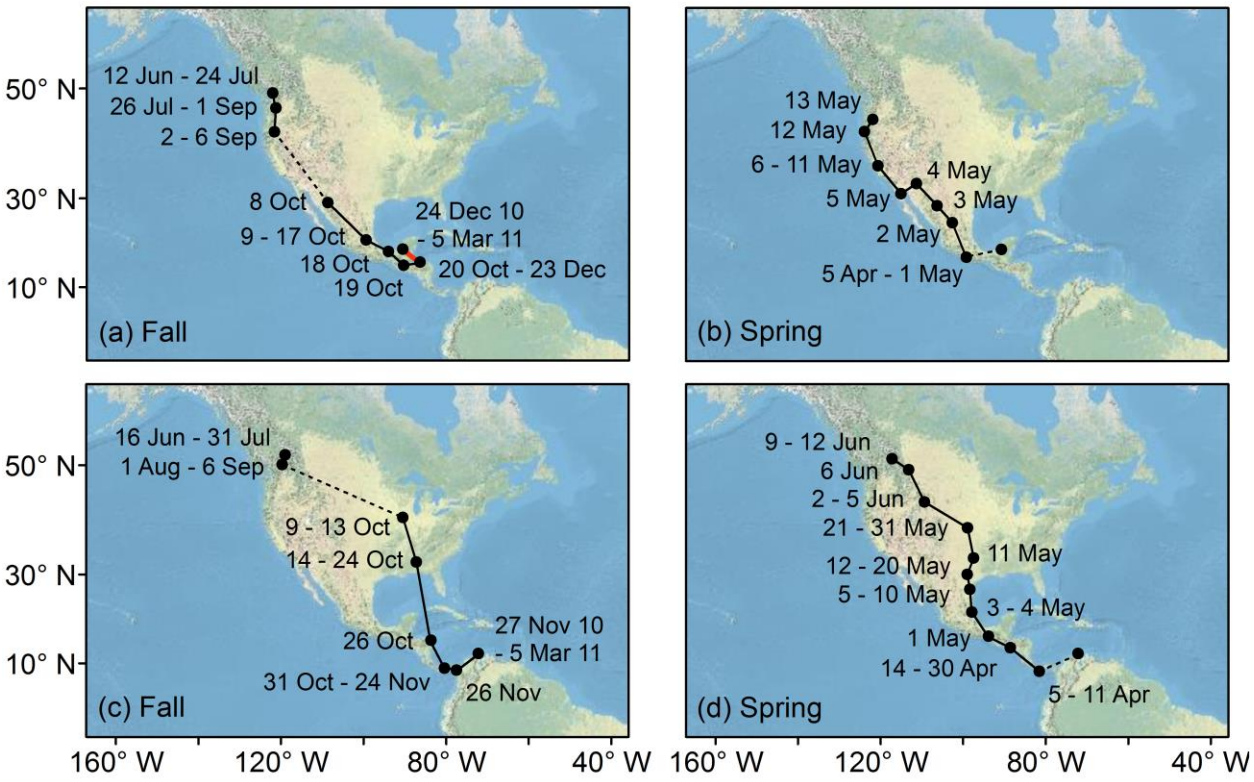


Figure 2.3 Detailed annual cycle of two Swainson's thrushes, one from the coastal subspecies group and one from the inland subspecies group. Individual from coastal group shown in panels a and b (shown in pink on Figure 2.1); individual from inland group shown in panels c and d (shown in red on Figure 2.1). Dates, stopover sites, breeding and wintering locations are shown. Fall migration for both birds shown in the left-hand panels; spring migration shown in the right-hand panels. Missing dates indicate periods for which the location of the bird could not be estimated (e.g., because of shading). Dashed lines link locations where latitude could not be estimated around the equinox periods. The bird in panels a and b employed two wintering sites; movement between these sites is shown using a thick red line.



Chapter 3: Hybrid songbirds employ intermediate routes in a migratory divide[2]

3.1 Introduction

Migratory divides are contact zones between related groups of organisms that breed in one place but use different routes to reach their wintering grounds. As migratory orientation of many organisms is partially genetically based (Berthold 1991, 1996; Helbig 1996; Pulido 2007; Liedvogel et al. 2011), an interesting question arises: if interbreeding occurs between groups with different migratory behaviours, what migratory behaviour will hybrids display? Helbig (1991) provided the clearest answer to this question to date, by using laboratory experiments to measure the instinctive migratory orientation of two forms of captive central European blackcap warblers and their hybrids: parental groups oriented southwest and southeast, whereas the hybrids oriented directly south. The parental orientations observed in the lab closely match the migratory routes of those populations in the wild, with the western group migrating southwest to Spain and western Africa, and the eastern group migrating southeast to the Middle East and eastern Africa. Helbig (1991) speculated that the orientation of hybrids would be an inferior behaviour in the wild, since it would tend to lead hybrids over the Alps and across the Mediterranean Sea and Sahara Desert. If migrating across such areas causes increased mortality or reduced condition, migratory behaviour could be a powerful source of selection against hybrids, thereby reducing gene flow between parental populations and promoting continued

[2] A version of this chapter has been published: Delmore KE, Irwin DE. 2014. Hybrid songbirds employ intermediate routes in a migratory divide. *Ecology Letters*, 17:1211.

divergence between the two forms (Helbig 1991; Irwin and Irwin 2005; Veen et al. 2007; Bensch et al. 2009; Rolshausen et al. 2009; Rohwer and Irwin 2011). We refer to this idea as the "migratory divide hypothesis."

Migratory divides have been documented in a wide variety of taxonomic groups (O'Corry-Crowe et al. 1997; Gagnaire et al. 2009; Altizer and Davis 2010), are especially common among songbirds (Irwin and Irwin 2005; Møller et al. 2011; Rohwer and Irwin 2011), and are highly relevant to the study of speciation (Helbig 1991; Irwin and Irwin 2005; Veen et al. 2007; Bensch et al. 2009; Rolshausen et al. 2009; Rohwer and Irwin 2011), phylogeographic history (Ruegg et al. 2006a), conservation and ecology (Boulet et al. 2006; Faaborg et al. 2010b). In many cases, a large area that is presumably difficult to migrate across (e.g., the Tibetan Plateau, deserts of the American Southwest, the Mediterranean Sea and Sahara Desert) separates the two migratory routes, suggesting inferiority of intermediate routes. Thus a variety of authors have postulated that migratory behaviour could play a major role in the speciation of birds and other taxa, yet only Helbig (1991, 1994) directly examined the orientation of hybrids, and those were captive birds assessed in the laboratory. To evaluate the migratory divide hypothesis most directly, freely-moving organisms need to be followed over the entire annual cycle.

We tracked birds in a divide between two subspecies groups of Swainson's thrushes: coastal, russet-backed (*Catharus ustulatus ustulatus*) and inland, olive-backed thrushes (*C. u. swainsoni*). These subspecies groups hybridize along the Coast Mountains of western North America and form a migratory divide, with the coastal group migrating along the west coast of North America to Mexico, Guatemala and Honduras and the inland group migrating through east-central North America to Columbia and Venezuela (Ruegg and Smith 2002; Delmore et al. 2012). Data from the hybrid zone between these groups suggests that they are at least partially

reproductively isolated from one another. Ruegg (2008) constructed clines for body size, plumage colour, mtDNA and AFLPs. These clines were concordant in width and narrow relative to what would be expected under neutral diffusion (i.e., no barrier to gene flow, observed width of 80 km and predicted width of 8000 km under neutral diffusion). Clines in these traits were also coincident across this hybrid zone and densities at its center were low (approximately 6 times lower than the edges of the zone, Ruegg 2008).

To track birds, we used light-level geolocators, archival tags that record light intensity at specific time intervals. These devices are attached to birds on the breeding grounds and retrieved the following year, when light intensity data are converted to day length and time of local midday and used to infer longitude and latitude over the entire year. Light-level geolocators were recently miniaturized (Stutchbury et al. 2009; McKinnon et al. 2013), permitting their use on songbirds including Swainson's thrushes. Specifically, we asked (1) whether routes of hybrids displayed intermediate or unusual characteristics compared to the parental forms, and (2) whether wintering sites of hybrids were different from those of parental forms.

Many of the hybrids tracked in our study employed intermediate migratory routes and wintering locations, supporting the first prediction from the migratory divide hypothesis. This hypothesis also predicts that intermediate routes and wintering areas employed by hybrids are inferior to those taken by parental forms. A direct test of this second prediction is beyond the scope of this paper. Nevertheless, previous work on this system suggested that the hybrid zone is too narrow to be explained purely by neutral diffusion, implying selection of some sort is acting against hybrids (Ruegg 2008). Using cline theory and molecular data collected along a transect through the hybrid zone we show that selection against hybrids is moderately strong in this

system. We suggest that the strikingly different and highly variable migratory behaviour of hybrids may be an important component of this selection.

3.2 Methods

3.2.1 Field sampling

We caught Swainson's thrushes using song playback and mist nets and banded them with Canadian Wildlife Service aluminum bands during the 2010 – 2013 breeding seasons. Photographs, blood samples, mantel feathers, and standard morphological measurements were obtained (blood samples from birds with geolocators were only taken on recapture to minimize stress on the birds when geolocators were applied).

3.2.2 Geolocators

We attached geolocators to birds at two sites (Pemberton and Hope, Figure 1.1) in the center of the hybrid zone in June 2011 (n = 58) and 2012 (n = 39). We used two models, the Mk20A (0.9 g, stalk length = 15 mm) and Mk10B (1.2 g, stalk length = 15 mm; British Antarctic Survey, Cambridge, United Kingdom). All of the birds tracked with geolocators were adult (AHY) males. Methods for the attachment and analysis of these devices are described in full in Delmore et al. (2012). Since we used two geolocators models in the present study, we calculated average Sun elevation angles (i.e., the angle of the sun above the horizon at which the light level indicates transition between night and day) for each model separately. Standard deviation (SD) among daily estimates of longitude and latitude while birds were on the breeding grounds was used to evaluate the accuracy of estimates from geolocators, as was the distance between the true deployment location and the location estimated using geolocator data. Similar to Delmore et al.

(2012) and other studies using this technology (Stutchbury et al. 2009), average SD for location estimates on the breeding grounds was 115 km for latitude and 73 km for longitude. Average Euclidean distance between the deployment location and estimates from geolocator data was 72 km. We also calculated estimates of location accuracy for each inferred stopover (defined as a location the bird used for more than 2 days) and wintering location, using the SD of daily estimates.

3.2.3 Genotyping

Populations at the center of the Swainson's thrush hybrid zone include both parental and hybrid individuals; previous research (Ruegg 2008) suggests that there is a full genetic continuum between the species in the hybrid zone, such that it likely has resulted from multiple generations of hybridization and backcrossing. We used three genetic markers fixed for differences between inland and coastal Swainson's thrushes to distinguish between parental genotypes and hybrid genotypes; eight markers were originally scanned (seven in nuclear introns and one mitochondrial sequence) and three showed fixed differences between the subspecies groups. We defined hybrid genotypes as those with a mixture of coastal and inland alleles at these markers.

We used a polymerase chain reaction-restriction fragment length polymorphism method to genotype birds at these markers. The mitochondrial marker is located in the mitochondrial control region and methods for genotyping birds at this marker are described in Ruegg and Smith (2002). Methods for genotyping birds at the remaining two nuclear introns are described in full in Appendix B.1. Briefly, we screened birds at two introns on the Z chromosome, one in the CHD1Z gene (Fridolfsson and Ellegren 1999) and one in the ADAMT56 gene (Backström et al. 2006). We amplified fragments from ten males (five from each taxon sampled in allopatry) and

sequenced these fragments bidirectionally. Males have two copies of the Z chromosome, and all analyses in this study were restricted to this sex. Diagnostic SNPs and restriction enzymes with diagnostic cut sites were identified in fragments from both genes; five birds in each taxon were homozygous for alternative alleles. The original sequences were obtained from blood samples of birds captured in Vancouver (allopatric coastal site) and tissue samples from the Field Museum of Chicago, from birds that were collected in Illinois and Wisconsin (allopatric inland sites). Birds at additional allopatric sites, including the Sunshine Coast and Vancouver (coastal sites, n = 19 for CHD1Z and 17 for ADAMT56) and William's Lake, Kamloops and Kelowna (inland sites, n = 36 for CHD1Z and 27 for ADAMT56) were homozygous for the expected allele.

3.2.4 Relationship between migratory behaviour and genetic background

We compared geolocator data from birds breeding in the hybrid zone (obtained in the present study) to geolocator data from birds breeding on either side of the hybrid zone (five coastal and four inland birds) collected by Delmore *et al*. (2012). The latter populations consisted of pure, parental individuals and are hereafter referred to as "allopatric".

We described the relationship between genetic background and both migratory routes and wintering sites qualitatively and quantified this relationship using linear models implemented in R 3.0.0 (R Development Core Team 2013). We measured genetic background as the number of inland alleles each bird had, migratory routes as the longitude at which birds crossed 30°N (and, in a supplemental analysis, 40°N) on fall and spring migration, and wintering sites as the longitude at which birds wintered. Birds were genotyped at two nuclear and one mitochondrial marker. Accordingly, they could have between 0 and 5 inland alleles (0 for pure coastal and 5 for pure inland). We considered locations when birds were stationary in the known wintering range

of Swainson's thrushes (Mack et al. 2000) as their wintering sites. We limited our analysis to longitude, which can be estimated with much less error than latitude using geolocator data. This is because latitude is estimated using day length (vs. time of local midday for longitude), which varies little across the low latitudes and also during the equinox. Longitude, by contrast, is estimated primarily using time of sunrise/sunset, which differs across longitudes at a wide range of latitudes (Ekstrom 2004; Fudickar et al. 2012; Lisovski et al. 2012). Five of the birds included in this analysis (two from the present study and three from Delmore et al. 2012) used two sites on the wintering grounds; we used the average longitude of these sites in our analysis. All linear models were performed both with and without allopatric individuals, and with and without parental genotypes in the hybrid zone.

3.2.5 Cline analyses

We estimated the strength of selection acting against hybrids using the formula $w = 4\sigma/(2s)^{1/2}$, where w is the width of the hybrid zone, σ is the organism's dispersal distance and s is the strength of selection acting against heterozygotes (Barton and Gale 1993). We used this equation with estimates of w and σ to estimate the strength of selection acting against hybrids. This method assumes that the cline is maintained by heterozygote disadvantage at a single locus and mating is random. Theoretical work has shown that this method holds for multilocus clines (Kruuk et al. 1999) and data from the Swainson's thrush hybrid zone suggests that mating is close to random (data from pairs suggest they do not mate assortatively and the hybrid zone is unimodal, K. Delmore unpub. data).

To obtain estimates of w, we captured and genotyped 134 birds at 9 sites between Vancouver and Kelowna (Appendix B.2). We collapsed these sites into a one-dimensional

transect by measuring the geographic distance between each site and Vancouver and fit sigmoidal clines to all three markers using the maximum likelihood approach implemented in CFit-7 (Gay et al. 2008). We treated each marker as a two-allele system, with each individual carrying one mitochondrial allele and two alleles from each nuclear marker. To estimate cline width we constrained all three clines to share the same value for width and center. We checked for convergence using different random seeds.

We used three estimates of σ in our analyses. One estimate came from the literature and is based on conservative estimates for dispersal in songbirds (σ = 10 km/generation; Brelsford and Irwin 2009). We estimated the remaining two values using linkage disequilibrium (D) and the equation $D = \sigma^2(1+r)/rw^2$, where r is the recombination rate between loci (Barton and Gale 1993). We estimated D between markers at sites within 20 km of the hybrid zones center and (1) between both nuclear markers and (2) between mitochondrial and nuclear markers. We calculated D between nuclear markers as $D = p_{AB} - p_A p_B$ using the genetics package in R, where p_{AB} = frequency of a gamete with upper-case alleles for loci A and B, p_A = frequency of allele A and p_B = frequency of allele B. We calculated D between mitochondrial and nuclear markers as $D = p_{MA} - p_M p_A$, where P_{MA} = frequency of a gamete with the mtDNA haplotype and the nuclear allele and p_M = frequency of allele M. Regarding the estimation of r, we note that both nuclear markers are on the Z chromosome. Nevertheless, they are distantly separated along that chromosome (in the zebra finch, the Z chromosome is ~80 Mb in length and these markers are located >25 Mb apart; genome.ucsc.edu, assembly July 2008). Combined with the observation that average recombination rate across the avian Z chromosome is 2.7 cM/Mb (Wahlberg et al. 2007), it is likely that 25 Mb is enough distance for these markers to behave as if they are not physically linked. Females only have one copy of the Z chromosome and, as a result,

recombination will not occur in this sex. Accordingly, we multiplied $r = 0.5$ by 2/3 to obtain the value 0.33 for the analysis between nuclear markers. We assumed a recombination rate of 0.5 in the analysis between mitochondrial and nuclear markers (i.e., not physically linked).

3.3 Results

We re-sighted 15/58 birds in June 2012 and 11/39 in June 2013. We successfully obtained geolocator data from all re-sighted birds in 2013 but were unable to obtain data from five in 2012 (two birds were re-sighted but not captured, two devices had fallen off, and we could not recover data from one device). Three of the ten birds recovered in 2012 had inland alleles at all genetic markers, one had coastal alleles, and six had a mixture of inland and coastal alleles. We refer to these genotypes as inland, coastal and hybrid. Two of the birds recovered in 2013 had inland genotypes, and the remaining nine had hybrid genotypes (Appendix B.3). Daily location estimates are shown in Appendices B.4 and B.5; the latter includes SD and standard error of latitude and longitude for days when birds appeared to be stationary; these standard errors on location estimates are generally very small compared to the differences among individuals in migratory routes and wintering locations.

3.3.1 Migratory behaviour is related to genetic background

Each allopatric group tracked in Delmore et al. (2012) and shown in black on Figures 3.1 and 3.2 showed high levels of migratory connectivity, following very similar routes and using the same long-term stopover sites on migration. The behaviour of birds genotyped as parental in our study was similar to birds in the former study (Figure 3.1 and Appendix B.6) and contrasts sharply with the behaviour of birds genotyped as hybrid; these birds exhibited tremendous variability in

the routes they employed, with many using novel routes that combined features characteristic of routes taken by inland and coastal allopatric birds. Some hybrids used routes that were geographically intermediate on fall and/or spring migration (e.g., Figure 3.2a orange, purple; Figure 3.2b red, orange, yellow). These routes took birds over broad arid and mountainous regions of the western United States that were not crossed by allopatric birds. Some hybrids used highly different routes southward and northward, taking routes similar to inland birds on fall migration (although somewhat further west in places), and coastal birds on spring migration (e.g., Figure 3.2ab dark blue, green, and purple; Figure 3.2cd dark blue and purple). These southward and northward routes were more than 2,000 km apart in some places (west to east). Finally, some hybrids used routes roughly similar to one or the other allopatric group (e.g., Figure 3.2cd pink, green, light blue and gray), although even these birds displayed periods of time when they were outside of the range of routes used by allopatric birds (e.g. light blue travelled further east in the United States during fall migration, and gray made a detour during fall migration from western Mexico northeast to the Gulf Coast).

Birds genotyped as hybrid in our study also exhibited high variability in their wintering sites. Inland birds tracked in Delmore et al. (2012) migrated to sites within 2 degrees longitude in Columbia and Venezuela; coastal birds migrated to sites within 8 degrees longitude between southern Mexico and Honduras (Figures 3.1 and 3.2). In contrast, birds tracked in the present study and genotyped as hybrid migrated to sites up to 18 degrees of longitude apart and to sites from southern Mexico through to Columbia (Figure 3.2).

Supporting the qualitative descriptions above, there is a strong relationship between genetic background (number of inland alleles) and migratory route (longitude at 30°N) on both fall and spring migration and wintering areas (Figure 3.3 solid lines). These relationships

remained strong after removing birds from allopatric sites (Figure 3.3 dashed lines and without filled points) and birds genotyped as parental (i.e., birds from both allopatric sites and hybrid sites with parental genotypes; Appendix B.7). We obtained similar results when these analyses were run using longitude at 40°N as our measure of migratory route (Appendix B.7).

Figure 3.4a shows the relationship between migratory routes on fall and spring migration; Figure 3.4b shows the relationship between spring migratory routes and wintering sites. Inland and coastal allopatric birds (squares) are clearly separated from one another on both axes in these panels. Hybrid zone birds (circles, gradient from black to white) show a wide variety of behaviours, with some clustering with allopatric birds, some similar to inland birds on fall migration and coastal birds on spring migration, and some intermediate to allopatric birds in one or both seasons. Overall, the genetic background of birds is strongly associated with behaviour in the hybrid zone, including the migratory routes and wintering sites chosen (Figures. 3.3 and 3.4).

Appendix B.3 summarizes data on the temporal and spatial patterns of migration for the birds tracked in this study and Delmore et al. (2012). On fall migration, birds with more inland alleles travelled longer distances at a faster pace (distance: $R^2 = 0.4$, $F_{1,28} = 18.39$, $P = 0.0002$; pace: $R^2 = 0.14$, $F_{1,28} = 4.36$, $P = 0.046$). On spring migration, birds with more inland alleles travelled longer distances and spent more time both on migration and at stopover sites (distance: $R^2 = 0.45$, $F_{1,27} = 22.11$, $P = <0.0001$; duration of migration: $R^2 = 0.14$, $F_{1,27} = 3.75$, $P = 0.06$; time spent at stopover sites: $R^2 = 0.14$, $F_{1,27} = 4.24$, $P = 0.049$).

3.3.2 Selection against hybrids is moderately strong

The cline estimated using genotypes of birds captured along a transect through the hybrid zone is 47 km in width. Using our estimate of dispersal distance from the literature ($\sigma = 10$

km/generation), we obtain an estimate for the strength of selection acting against heterozygotes (*s*) of 0.36. The alternative method (which uses linkage disequilibrium and does not require an estimate of dispersal distance) leads to estimates of linkage disequilibrium (*D*) of 0.2 between nuclear markers and 0.07 between mitochondrial and nuclear markers. These values for *D* give estimates of 10.47 and 7.18 km/generation for σ and estimates of 0.40 and 0.19 for *s*. Together, these results suggest that the strength of selection acting against heterozygotes (*s*) as measured at the loci under study lies roughly between 0.19 and 0.40.

3.4 Discussion

This is the first study to track freely-moving organisms from the center of a migratory divide over the entire annual cycle. Compared to parental forms from both allopatric and hybrid sites, hybrids exhibited tremendous variability in the migratory routes they employed. A large number of them used intermediate or mixed routes on migration. The mixed routes were a novel and unexpected finding that can also be viewed as a form of intermediacy. Hybrids also exhibited much variability in the wintering sites they chose, which spanned almost the entire range of longitudes between parental wintering sites. Our study builds on Helbig's (1991) study of orientation in captive, lab-raised hybrids by tracking free-flying birds from a natural hybrid zone across the entire annual cycle, including both migratory seasons and the winter.

Previous work in the Swainson's thrush hybrid zone suggested that hybrids are selected against (Ruegg 2008). We used cline theory to estimate the strength of this selection. The values we calculated were high (s = 0.19 – 0.40) compared to other hybrid zones (toads, s = 0.17 – 0.22 Szymura and Barton 1986; grasshoppers, s = 0.06 – 0.4 Kawakami et al. 2008; warblers, s = 0.08 – 0.28 Brelsford et al. 2009), suggesting that selection is moderately strong. Inferences from the

birds tracked in our study suggest that differences in migratory orientation could contribute to this selection. Specifically, many of the hybrids that used intermediate routes on migration navigated over mountain ranges and arid regions in the western United States. These areas are likely difficult to navigate over and provide limited habitat to forage, especially during fall migration when conditions in western North America are dry and unproductive, forcing many species to migrate south before undergoing their annual molt (Rohwer and Irwin 2011). An examination of the Normalized Differential Vegetation Index (NDVI) for this region supports this suggestion; NDVI reflects vegetation conditions and is much lower in the area between the routes employed by allopatric populations (Figure 3.5). If hybrids suffer increased mortality or reduced condition after migrating over these areas they will likely be selected against, reducing gene flow between subspecies groups of Swainson's thrushes and promoting continued divergence.

Two aspects of our data cast some doubt on the supposition that intermediate routes contribute to selection against hybrids. First, we recovered 15 birds with hybrid genotypes, showing that many hybrids are able to migrate and return successfully to the breeding grounds. Second, the return rates for birds tracked with geolocators were similar in allopatric and hybrid populations (return rate, 95% confidence interval: coastal = 25%, 11 – 47; inland = 26%, 11 – 49; hybrid = 27%, 19 – 36). Nevertheless, the confidence intervals around these estimates are quite large, leaving the possibility of substantial differences in survival. In addition, migration can influence fitness through effects on reproduction rather than survival. For example, condition and spring departure date have been shown to affect reproductive success in American redstarts (Marra et al. 1998). Moderately reduced reproduction (e.g., roughly 20-40% reduced) because of migratory routes would be sufficient to account for the estimated strength of selection against

hybrids described above. Finally, we note that coastal and inland Swainson's thrushes differ in several traits, including song, morphometrics, colour and migratory routes, but differences in migratory routes distinguish these groups much more clearly than the other traits (Cohen's *D* of 9.01 for migration, 3.61 for song, 2.15 for morphometrics and 0.28 for colour; Appendix B.8). Traits that best distinguish forms are more likely to play a strong role in selection against hybrids and the generation of reproductive isolation (Safran et al. 2012). Altogether, we feel it is likely that the intermediate routes of hybrids are inferior, but further research is required to fully test this hypothesis.

The mixed strategy employed by some of the hybrids tracked in our study was unexpected and could serve as an additional source of selection against hybrids. Specifically, naïve first time migrants rely primarily on vector navigation (i.e., time and direction) to reach their wintering grounds. Vector navigation is believed to be largely endogenous and genetically determined (Berthold 1991, 1996). Migrants are thought to develop navigational maps from cues they learn on the first leg of migration to help complete subsequent trips. The development of these maps is beneficial, as they allow migrants to take more direct routes, revisit good stopover sites and avoid unfavourable areas (Berthold 1996; Wiltschko and Wiltschko 1999). Juvenile hybrids using the mixed strategy will not benefit from any experience gained on their first outward journey, as they will return along a different path, and could experience fitness cost as a result. Nevertheless, this cost may not be large, as there is some innate control of return migration (Wiltschko and Wiltschko 1999) and much of the information gained by first time migrants comes from their post-fledgling period when they build maps of their home area (Wiltschko and Wiltschko 1999). It is also common for birds to take more direct routes on spring migration (Berthold 1996; Wiltschko and Wiltschko 1999). Interestingly, it is also possible that

hybrids using the mixed strategy avoid fitness costs associated with taking a purely intermediate route on fall and/or spring migration, as they are following routes used by parental forms and around unsuitable areas on each leg of migration.

Our results also allow tentative inferences about the control of bird migration. Specifically, the mixed strategy employed by a few of the hybrids we tracked suggests that orientation on fall and spring migration may be regulated differentially. This argument has been made by authors who have documented examples of loop migration, where birds use different routes to navigate around geographic features on fall and spring migration (e.g., red-backed shrikes and the Arabian Peninsula, Tøttrup et al. 2012) and has been shown for the initiation of migration, with gonadal development and hormones playing a role in the initiation of spring migration but not fall migration (Berthold 1996). We also documented strong relationships between the genetic ancestry of individuals and both their migratory routes and wintering locations. Combined with the observations that hybrids took intermediate routes and parental forms from the center of the hybrid zone followed routes more similar to allopatric populations (including the use of similar long-term stopover sites, intratropical migration by the bird genotyped as coastal and loop migration by the birds genotyped as inland; Appendix B.6), these findings suggest that migratory orientation is genetically controlled in Swainson's thrushes and additively inherited. This suggestion supports results from the crossbreeding experiments conducted by Helbig (1991, 1994, 1996) but is not in accordance with Veen et al. (2007). These authors used stable isotopes to infer the wintering grounds of collared and pied flycatchers that form a divide in central Europe and found that hybrids had similar isotopic values to pure, pied flycatchers. These results suggest that at least for genes controlling wintering locations, the genes

of pied flycatchers are dominant to those of collared flycatchers. It is also possible that hybrids and pied flycatchers winter in different areas that have similar isotopic signatures.

The three genetic markers used in our study sample a very small proportion of the genome, are highly unlikely to code for differences in migratory orientation, and allow only a rough estimation of ancestry (e.g., they preclude the use of programs like STRUCTURE to assign ancestry estimates, Pritchard et al. 2000). Despite the noise in our ancestry estimates, we uncovered a strong relationship between genetic ancestry and migratory behaviour, suggesting that future genetic analyses including additional genetic markers would show an even stronger relationship. It would also be instructive to follow the same birds over multiple years, providing further evidence for the genetic basis of this trait (if the trait is largely genetically determined birds should follow roughly the same route each year, Stanley et al. 2012).

Light-level geolocators record light intensity on scale between 0 and 60. These data are used along with a chosen threshold light intensity level to determine the time of sunrise and sunset. Calibration data (i.e., data from devices when they are at a known location) are then used to determine the sun's elevation angle (or solar zenith) at the chosen threshold level. This variable is used along with the time of sunrise and sunset to infer daily longitude (calculated as the average of sunrise and sunset, time of local noon) and latitude (calculated as the difference between sunrise and sunset times, day length; Ekstrom 2004). This technology is associated with a few limitations. For example, there is little variation in day length near the equinoxes and the equator, making it difficult to estimate latitude at these times. Geolocators are also sensitive to shading, as these events affect the relationship between the sun’s elevation angle and the chosen threshold level. For instance, on a cloudy day or in a closed habitat the threshold will be met when the sun is at a higher angle (Ekstrom 2004; Fudickar et al. 2012; Lisovski et al. 2012).

Despite these limitations, data from stationary geolocators and geolocators attached to resident taxa (i.e., birds that are in a known location throughout the year) have shown that estimates from these devices are still reliable, especially when steps are taken to reduce sources of error (Fudickar et al. 2012; Lisovski et al. 2012). These steps include choosing a threshold level deeper into twilight (i.e., where variation in light is steeper; Ekstrom 2004) and using in-habitat calibration data (i.e., data from geolocators deployed in the known habitat; Lisovski et al. 2012). Authors have also advocated reliance on estimates of longitude, as they are less affected by shading (Ekstrom 2004; Stapley et al. 2010a; Fudickar et al. 2012; Lisovski et al. 2012). For example, Ekstrom (2004) showed that in partly cloudy weather sunrise and sunset transitions are reasonably symmetrical and, as a result, will have less of an effect on longitude as this variable is estimated by taking the average of sunrise and sunset. This will not be the case for latitude, which is estimated as the difference between sunrise and sunset. We followed all of these suggestions in our study, using a threshold level of 1, calibration data from devices deployed on birds during the breeding season and relying heavily on estimates of longitude. We note, that even with these limitations geolocators remain the best technology available to track long distance migrants less than 50 g; satellite transmitters and GPS loggers are still limited to large animals, band recovery data is often associated with small sample sizes and biological markers (e.g., stable isotopes) only provide indirect information on migratory routes (Webster et al. 2002). In addition, we note that the error in longitude we reported (average SD on breeding: 73 km or 1.02 degrees; average SE on breeding: 12 km or 0.16 degrees) is far less than the difference observed between coastal and inland routes (between 15 and 20 degrees).

Long-distance migration is undertaken by individuals from nearly every major animal group and has been the focus of research for decades. Despite this ubiquity and fascination, we

still know very little about how this behaviour is accomplished and its contribution to speciation (Alerstam 2006; Pulido 2007). In this study, we used the power of a modern tracking technology to delve further into these topics. We present evidence that differences in migratory orientation could play an important role in divergence between subspecies groups of Swainson's thrushes. This provides a rare demonstration of the possible role of a behavioural trait in speciation; this phenomenon could be particularly important in birds, as migratory divides appear to be relatively common in this group (Irwin and Irwin 2005; Møller et al. 2011; Rohwer and Irwin 2011) and several reviews have suggested that traits reducing gene flow following mating (i.e. postmating barriers) could be central to speciation in this group (Grant and Grant 1992; Price and Bouvier 2002; Rabosky and Matute 2013). These results highlight the power of modern tracking techniques for studying long-distance migration and the utility of hybrid zones for studying traits involved in speciation.

Figure 3.1 Annual cycle of Swainson's thrushes genotyped as parental from the hybrid zone. Data for birds recaptured in 2012 (a and b) and 2013 (c and d) are shown. Routes for birds recaptured in 2012 and genotyped as inland are shown in light blue, dark green and white (a and b); the bird genotyped as coastal and recaptured the same year is shown in pink (a and b). Routes for birds recaptured in 2013 and genotyped as inland are shown in maroon and white (c and d). Black lines show routes taken by allopatric birds from Delmore et al. (2012; four coastal birds and five inland birds). Dashed lines link locations where latitude could not be estimated around the equinox periods.

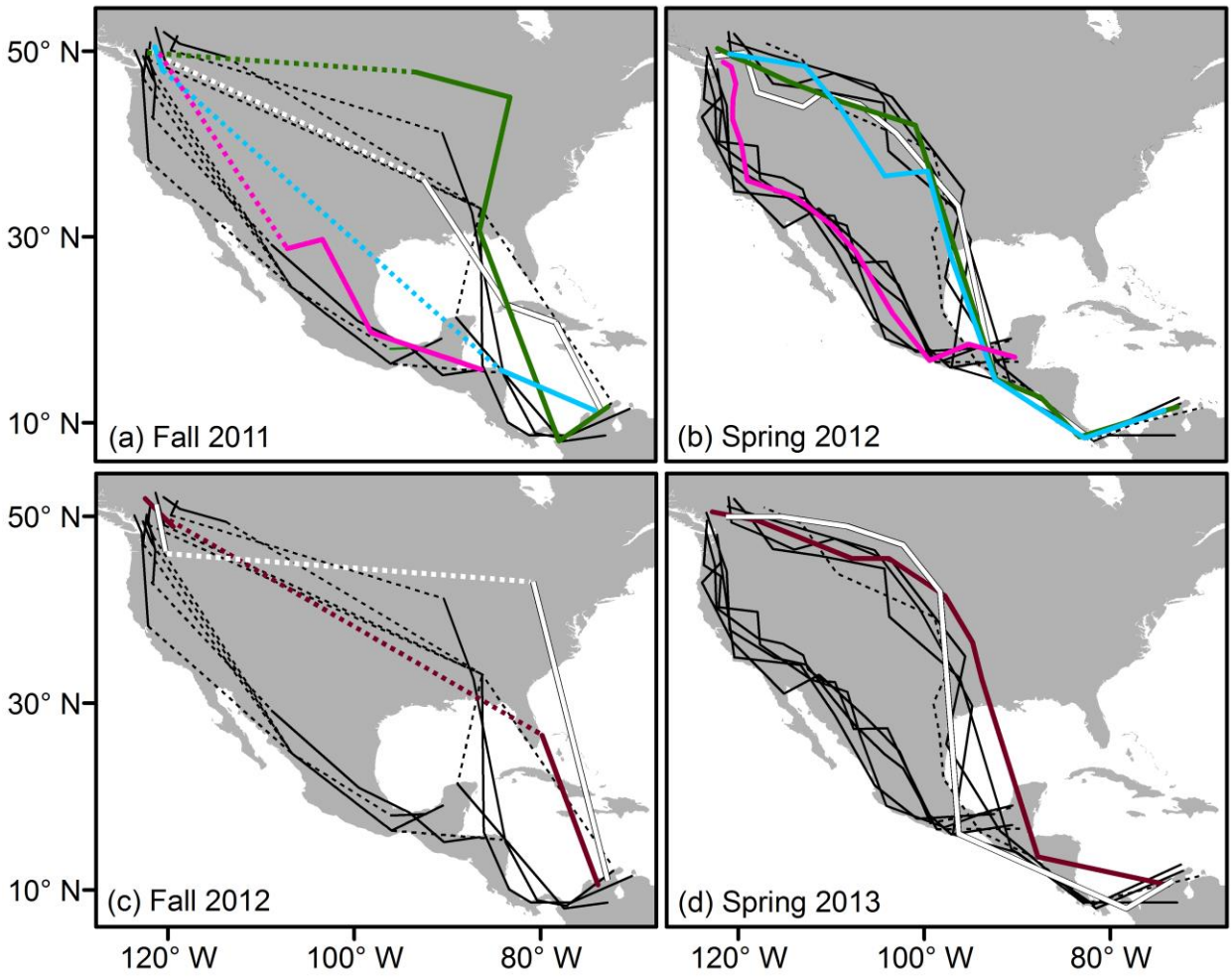


Figure 3.2 Annual cycle of hybrid Swainson's thrushes. One set of birds was tracked between 2011 and 2012 (ab) and a second separate set between 2012 and 2013 (cd). Fall (ac) and spring (bd) migration are shown. Black lines show routes taken by allopatric birds from Delmore et al. (2012; four coastal birds and five inland birds). Dashed lines indicate periods near the equinox, when locations could not be determined accurately due to little latitudinal variation in day length.

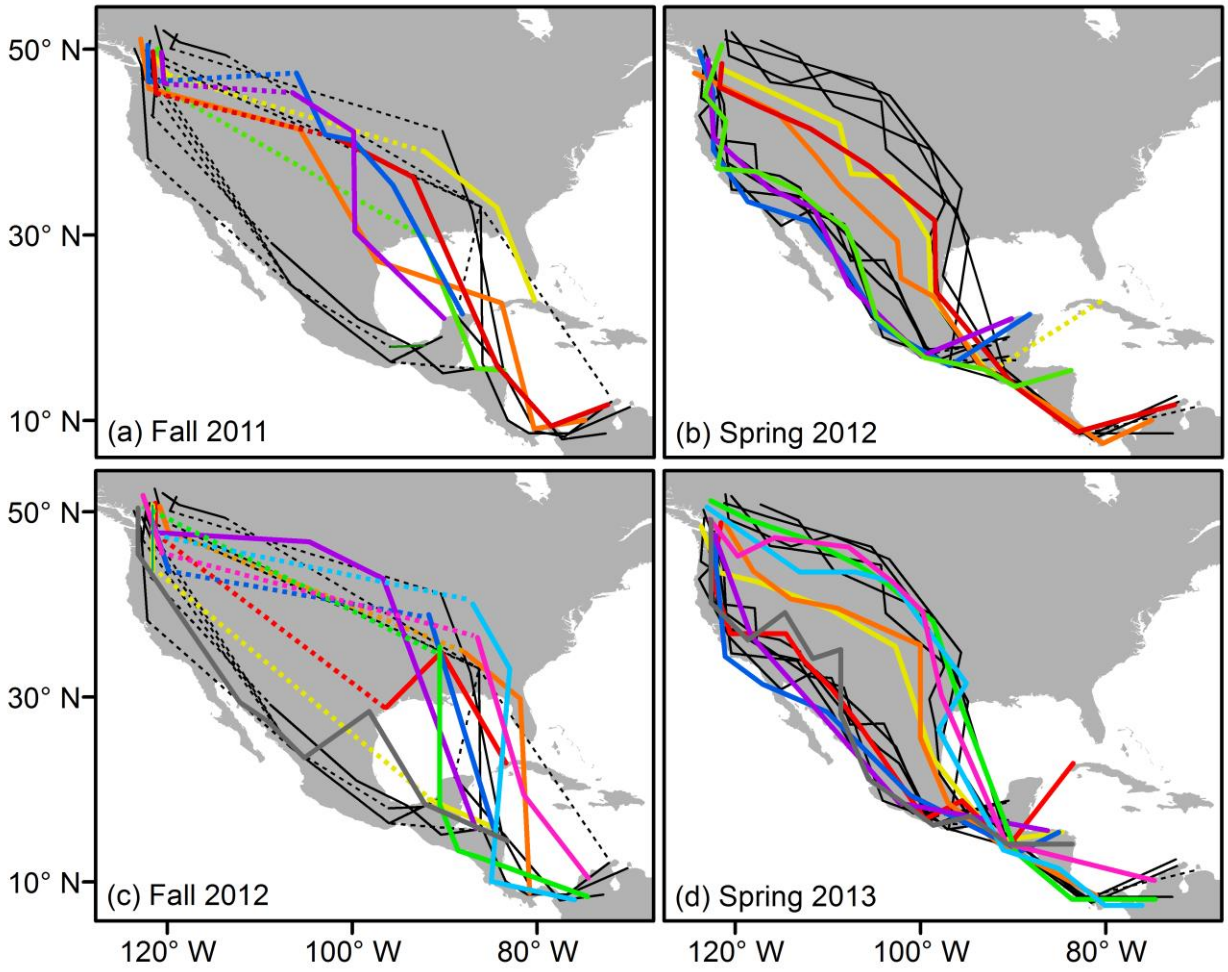


Figure 3.3 Relationship between genetic background and longitude during the annual cycle. Genetic background was quantified as the number of inland alleles an individual has and is correlated with longitude (°W) at 30°N latitude on (a) fall migration, (b) spring migration and (c) the wintering grounds. Birds from the hybrid zone are shown with open circles; birds from allopatric sites are shown with filled circles. Results from linear models are shown for analyses including all the data (solid lines, fall migration: $R^2 = 0.61$, $F_{1,28} = 44.51$, $P < 0.0001$; spring migration: $R^2 = 0.66$, $F_{1,28} = 54.18$, $P < 0.0001$; wintering grounds: $R^2 = 0.77$, $F_{1,28} = 92.09$, $P < 0.0001$) and excluding allopatric birds (dashed line, fall migration: $R^2 = 0.36$, $F_{1,19} = 10.56$, $P = 0.004$; spring migration: $R^2 = 0.54$, $F_{1,19} = 22.61$, $P = 0.0001$; wintering grounds: $R^2 = 0.57$, $F_{1,19} = 25.39$, $P < 0.0001$).

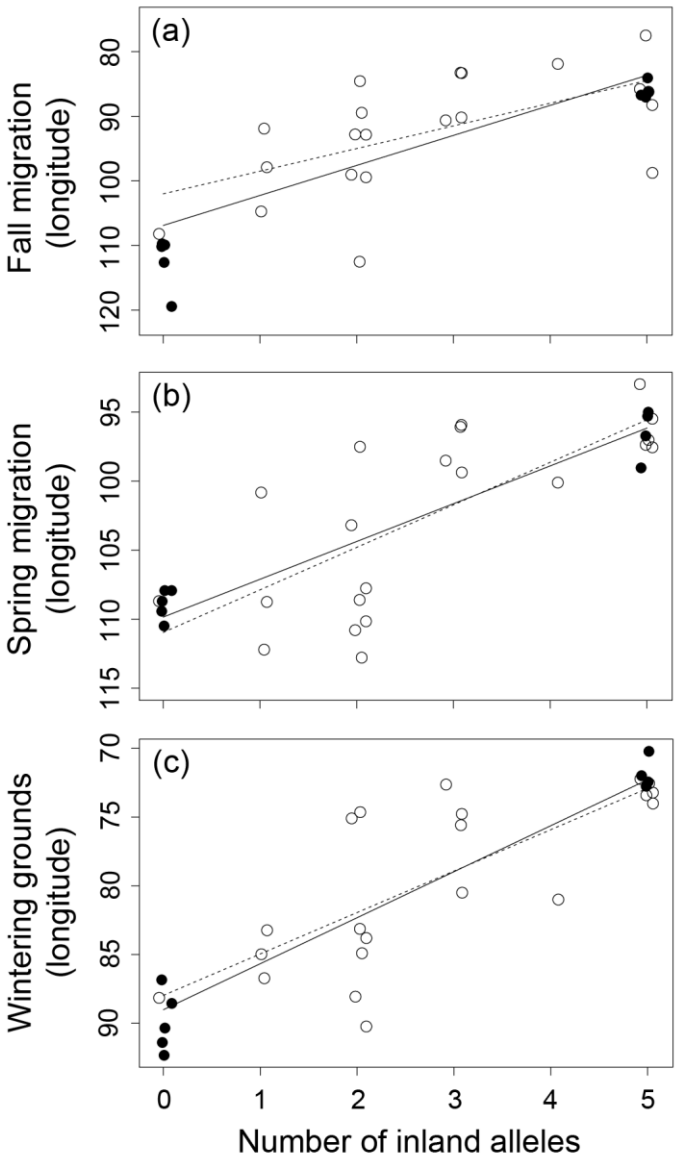

(a)
Fall migration (longitude)
120
110
100
90
80
(b)
Spring migration (longitude)
115
110
105
100
95
(c)
Wintering grounds (longitude)
90
85
80
75
70
0
1
2
3
4
5
Number of inland alleles

Figure 3.4 Relationship between longitude on spring migration, fall migration and the wintering grounds. Longitude (°W) on migration was measured at 30°N latitude. Data from allopatric birds are shown with squares (black = coastal, white = inland). Data from birds in the hybrid zone are shown with circles and varying colours from black to white (black = 0 inland alleles, grays decreasing in darkness = 1 – 4 inland alleles, white = 5 inland alleles).

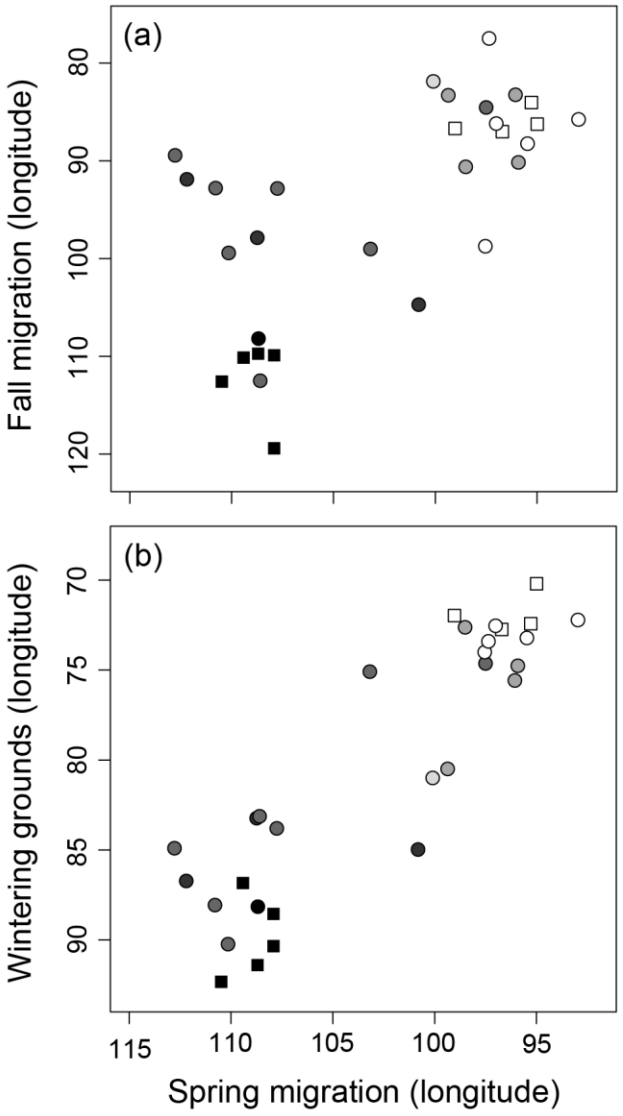
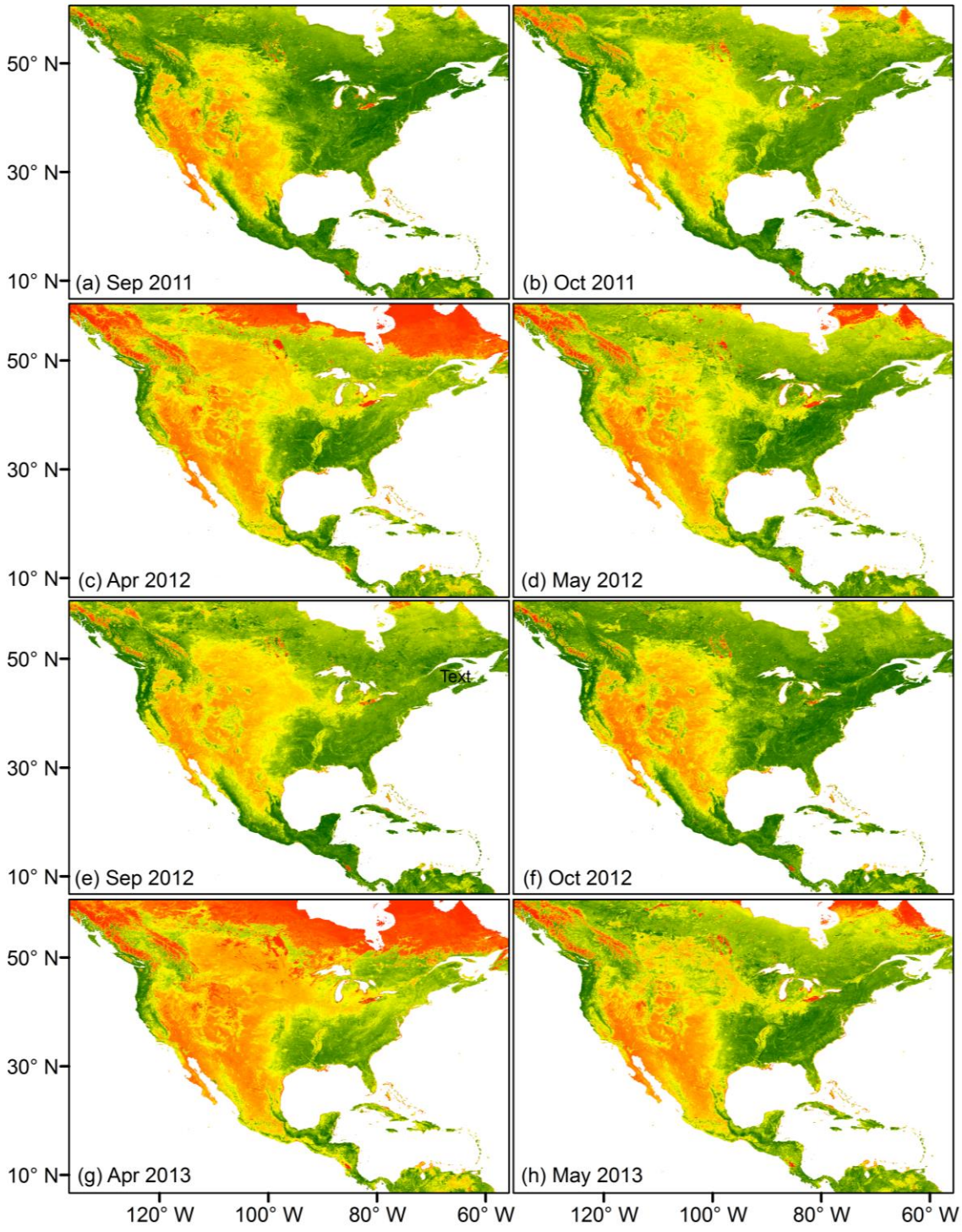


Figure 3.5 MODIS/Terra Normalized Differential Vegetation Index (NDVI) from the migratory periods included in this study. Index was downloaded as a 0.05° resolution from the Land Processes Distributed Active Archive Center. Data for both fall (Sep – Oct) and spring (Apr – May) migration and for both years of tracking (2011 – 2012 and 2012 – 2013) are shown. More green values reflect healthier vegetation (more reflectance in the near infrared part of the spectrum, lpdaac.usgs.gov).


50° N
30° N
10° N
(a) Sep 2011
(b) Oct 2011
(c) Apr 2012
(d) May 2012
Text
(e) Sep 2012
(f) Oct 2012
(g) Apr 2013
(h) May 2013
120° W
100° W
80° W
60° W

Chapter 4: Genomic analysis of a migratory divide reveals candidate genes for migration and implicates selective sweeps in generating islands of differentiation[3]

4.1 Introduction

The process of speciation often begins with a period of geographic isolation. During this time, gene flow between populations is restricted and divergent selection on traits can lead directly or indirectly to reproductive isolation (Coyne and Orr 2004; Price 2008). Sources of divergent selection and reproductive isolation are well described in many groups. Among songbirds, for example, differences in habitat choice, song and seasonal migration likely help maintain taxonomic boundaries (Price 2008; e.g., Haavie et al. 2004; Bearhop et al. 2005; Smith and Benkman 2007). Research on the genetic basis of divergent selection and reproductive isolation has been conducted since the early 20th century (e.g., Fisher's [1930] mathematical model highlighting the role mutations of small effect have in adaptation, Dobzhansky's [1936] genetic mapping experiments with lab crosses, and Barton and Hewitt's [1985, 1989] work using clines to estimate the number of genes generating barriers to gene flow). Nevertheless, questions on this topic still remain, and include: how many regions of the genome are affected by selection, and what is their size, location and gene content (Orr 2005; Noor and Feder 2006; Stapley et al. 2010b; Nosil and Feder 2012)? The development of next generation sequencing techniques has facilitated research on this topic, permitting the assembly of reference genomes for non-model

[3] A version of this chapter has been published: Delmore KE, Hübner S, Kane NC, Schuster R, Andrew RL, Câmara F, Guigo Roderic, Irwin DE. 2015. Genomic analysis of a migratory divide reveals candidate genes for migration and implicates selective sweeps in generating islands of differentiation. *Molecular Ecology*, 24:1873.

organisms and examinations of differentiation on a genomic scale (Metzker 2010; Davey et al. 2011).

A growing number of studies have documented genome-wide heterogeneity in estimates of genetic differentiation (Ellegren et al. 2012; Nadeau et al. 2013; Renaut et al. 2013). Two main models have been proposed to explain these patterns. First, areas of increased genetic differentiation ('islands of differentiation') could harbour one or more loci involved in maintaining reproductive isolation (Wu 2001; Nosil et al. 2009; Feder et al. 2012; Via 2012). These regions will be unaffected by gene flow following hybridization and introgression, while regions of the genome that do not harbour these loci will be homogenized by gene flow following these processes (Wu 2001; Nosil et al. 2009; Feder et al. 2012; Via 2012). A second model was described by Noor and Bennett (2009) and expanded on by Hahn and colleagues (Turner and Hahn 2010; Cruickshank and Hahn 2014). These authors argued that differences in the strength of selection alone could generate genomic heterogeneity in differentiation, even in complete allopatry. They noted that genetic distance is often quantified using F_{ST}, which is a relative measure of differentiation that can be elevated by not only increased absolute differentiation between populations but also reductions in within population diversity (Cruickshank and Hahn 2014). Several processes can reduce within population diversity, including selective sweeps at advantageous mutations that also reduce variation at linked neutral sites (Maynard Smith and Haigh 1974). Under this model, regions of the genome experiencing low or no linked selection are predicted to exhibit lower levels of relative differentiation (Cruickshank and Hahn 2014). These two models make differing predictions regarding patterns of absolute differentiation between populations: in the divergence-with-gene-flow model absolute differentiation should be high in regions of high relative differentiation, whereas in the

selection-in-allopatry model absolute differentiation should be roughly constant throughout the genome, as variation in this parameter will be driven primarily by the accumulation of nearly-neutral mutations rather than selectively advantageous mutations (due to their relative frequency, Nosil et al. 2009; Cruickshank and Hahn 2014) .

Features of genomes that lead to low polymorphism may also contribute to the generation of genomic heterogeneity. For example, following from the selection-in-allopatry model referred to above, linked selection and/or lower mutation rates in areas of reduced recombination like centromeres are expected to maintain lower within population levels of diversity in these areas, leading to increased levels of relative differentiation (Kulathinal et al. 2008; Noor and Bennett 2009; Cruickshank and Hahn 2014). Sex chromosomes also often exhibit higher rates of relative divergence. This association can be at least partially attributed to reduced rates of recombination, as recombination will only occur in the homogametic sex for the common sex chromosome and not at all for the rare chromosome (e.g., Z vs. W in systems with ZW-sex determination). Several other factors could also account for this association, including higher densities of genes associated with reproductive isolation on the sex chromosomes (Mank et al. 2007; Presgraves 2008) and Haldane's rule (Haldane 1922). Haldane's rule posits that when hybrids are formed the heterogametic sex will experience greater fitness losses. Genes that reduce hybrid fitness are often recessive and interactions between these genes are often implicated in reproductive isolation (Coyne and Orr 2004). Because genes on the sex chromosomes will always be expressed in the heterogametic sex, they will have a greater effect than autosomal genes.

Here we examine patterns of genomic differentiation between coastal and inland subspecies groups of the Swainson's thrush (*Catharus ustulatus ustulatus* and *C. u. swainsoni*). These groups breed throughout much of Canada and the northern United States and form a

narrow hybrid zone along the Coast Mountains (Ruegg 2008). This hybrid zone is characterized by coincident clines in morphological and genetic traits, low population densities and moderate selection against hybrids (Ruegg and Smith 2002; Ruegg 2008; Delmore and Irwin 2014). Paleodistribution models and phylogenetic analyses suggest that these groups diverged during the Pleistocene, when their ranges were restricted to separate glacial refugia (Ruegg and Smith 2002; Ruegg et al. 2006a). Potential sources of reproductive isolation between these groups are well described, and include differences in their climatic niches (Ruegg et al. 2006a), plumage (Ruegg 2008), song (Ruegg et al. 2006b) and migratory behaviour (Ruegg and Smith 2002; Delmore et al. 2012; Delmore and Irwin 2014).

Differences in migratory behaviour distinguish Swainson's thrushes most clearly (Delmore and Irwin 2014), with the coastal group migrating south to Central America and the inland group southeast to South America (Ruegg and Smith 2002; Delmore et al. 2012). Contact zones between taxa with different migratory routes have been termed 'migratory divides', and there is considerable interest in the role these divides could play in speciation (Helbig 1991; Bearhop et al. 2005; Irwin and Irwin 2005; Bensch et al. 2009; Rolshausen et al. 2009; Rohwer and Irwin 2011). The Swainson's thrush is the only system where hybrids have been tracked over the entire annual cycle. Hybrids used intermediate routes to parental forms on migration that took them over several geographic barriers in the southwestern United States, likely reducing their fitness (Delmore and Irwin 2014).

A number of genetic analyses have been conducted in the Swainson's thrush system. For example, Ruegg and Smith (2002) sequenced the control region of mitochondrial DNA from birds across the breeding range, identifying five diagnostic mutations between the coastal and inland groups and estimating a net sequence divergence of 0.69%. More recently, Ruegg et al.

(2014) used restriction site-associated DNA (RAD) sequencing to describe patterns of genomic differentiation between populations at extreme ends of the species range. These authors documented a heterogeneous pattern of relative differentiation, with elevated F_{ST} both on the Z chromosome and near inferred locations of centromeres. Candidate genes thought to be associated with traits involved in reproductive isolation (song, plumage and migration) did not significantly cluster in islands of differentiation, although genes associated with migration did have more high F_{ST} SNPs than expected by chance. Ruegg et al. (2014) interpreted these findings as showing that features of the genome and the retention of ancestral polymorphisms were more important for generating genomic heterogeneity than gene flow. These authors relied on the zebra finch genome to position SNPs and identify candidate genes.

We aim to compliment and expand on previous genetic surveys here, by assembling and annotating a draft reference genome for the Swainson's thrush and generating whole genome shotgun (WGS) sequence data for populations directly adjacent to the hybrid zone. We quantified relative genetic differentiation using F_{ST} to confirm previous findings of heterogeneity in this parameter across the genome. We then used additional population genetic parameters to provide more detailed insight into the processes that generate this pattern. Specifically, we estimated absolute genetic differentiation between inland and coastal Swainson's thrushes using d_{XY}, which is the average number of pairwise differences between sequences of two populations and does not depend on variation within populations (Nei 1987; Cruickshank and Hahn 2014). We also estimated the level and structure of diversity within coastal and inland populations using π (average number of pairwise differences between sequences of different individuals within a population) and Tajima's D (difference between the total number of segregating sites and π). If variable gene flow generates genomic heterogeneity, absolute genetic differentiation should be

elevated in areas of high F_{ST}, because heterogeneity results from the introgression of alleles between the forms in the low-F_{ST} regions. Conversely, if variable selection in allopatry has generated heterogeneity, regions of high F_{ST} should have low within population diversity and an excess of low frequency polymorphisms (i.e., negative values of Tajima's *D*). Absolute genetic differentiation should also be relatively constant, since heterogeneity results from selective sweeps in a background of nearly-neutral mutations under this model. We examined levels of differentiation near centromeres and on the Z chromosome to implicate features of the genome in generating genomic heterogeneity.

The use of WGS sequencing and an annotated reference genome greatly increased genomic resolution compared to previous genetic surveys, as more reads mapped to our species-specific reference and WGS sequencing identified SNPs located throughout the genome (not just near restriction enzyme cut sites). We used this increased resolution to re-evaluate one particular result emphasized by Ruegg et al. (2014), that candidate genes for migration were not clustered in islands of relative differentiation. These authors included a list of 25 candidate genes but only had coverage of five. The greater coverage of genes in our analysis permits a more thorough test of a possible association. We looked for candidates in the list compiled by Ruegg et al. (2014) and generated a second list of candidates that included genes functionally related to those linked to migration. Several recent studies have improved our understanding of migration's genetic basis (Jones et al. 2008; Steinmeyer et al. 2009; Mueller et al. 2011), but given the complexity of this trait it is likely that many genes have yet to be identified. Ruegg et al. (2014) also used a relatively large window size (1000 kb) to identify islands; the greater density of SNPs in our study enabled us to reduce this window size an order of magnitude and examine more fine scale patterns. This work informs questions concerning the genetic architecture of reproductive

isolation using a well-characterized avian system. We generate useful resources for future analyses (e.g., the first reference genome assembled for a Nearctic-Neotropical migrant) and provide a comparison between different sequencing technologies and approaches.

4.2 Methods

4.2.1 Draft reference genome

We assembled a draft reference genome for the Swainson's thrush using DNA from an adult male captured in the range of the coastal subspecies group (Vancouver; Figure 1.1; caught June 1 2011, ID# KF01K01). This individual was caught using song playback and a mist net. Standard morphological measurements and photographs were taken, along with a blood sample. Males are the homogametic sex in birds and, as a result, we have no coverage of the W chromosome but do have similar coverage of the Z and autosomal chromosomes.

We extracted DNA from the blood sample using standard phenol-chloroform methods and had three WGS libraries constructed and sequenced: one fragment library (180 bp insert) at the Biodiversity Research Center at the University of British Columbia and two mate pair libraries (2 and 4 kb inserts) at Macrogen in Korea. All libraries were sequenced on Illumina HiSeq2000 machines with paired-end, 100 bp sequencing technology. The fragment library was sequenced in one lane; the mate pair libraries were uniquely barcoded and sequenced together in one lane. Methods for trimming, assembling and annotating the reference are in Appendix C.1. Briefly, custom perl scripts were used to trim libraries, which were then aligned to the zebra finch genome using Stampy-1.0.22 (Lunter and Goodson 2011) and built into a consensus sequence using *samtools* (Li et al. 2009). The genome was annotated using the MAKER pipeline (Cantarel et al. 2008).

4.2.2 Population samples, sequencing and mapping

Two transects across this hybrid zone have been the focus of previous studies (Ruegg 2008; Delmore and Irwin 2014). We collected blood samples from populations at both ends of each of these transects (n = 10 male birds/population; 4 populations; Figure 1.1), using the same methods described above to capture birds and extract DNA. We followed methods described in Sham et al. (2002) to pool DNA from each population; we quantified the DNA using a Qubit fluorometer (Life Technologies) and diluted each sample to 25 ng/μL. Before pooling DNA, we amplified two genes in each sample to ensure successful amplification of DNA from all individuals (CHD1Z and ADAMT56, methods described in Delmore and Irwin 2014).

Small insert (300 bp) WGS libraries were constructed for each pool at the Biodiversity Research Center at the University of British Columbia. All libraries were sequenced on an Illumina HiSeq2000 with paired-end, 100 bp sequencing technology. Each pool was uniquely barcoded and run with one other pool in a single lane (barcodes provided in Appendix C.2). Methods for trimming and aligning these data to the reference are provided in Appendix C.3.

4.2.3 Estimation of population genetic parameters

Following alignment, we combined data from pools of the same subspecies group and estimated four population genetic parameters: F_{ST}, d_{XY}, π and Tajima's D. We used Popoolation2 (Kofler et al. 2011b) to estimate F_{ST}. Popoolation2 estimates F_{ST} as: $(H_T - H_S)/H_T$, where $H = (C/(C-1))\ (1 - f_A^2 - f_a^2)$, C = coverage of pool, f_N = frequency of nucleotide N, H_S = average heterozygosity within subpopulations and H_T = heterozygosity in total population. We estimated d_{XY} at all sites between populations as: $\left(f_{A,x}f_{a,y}\right) + (f_{A,y}f_{a,x})$, where x and y designate different

populations and used Popoolation (Kofler et al. 2011a) to estimate both π and Tajima's *D*. All four parameters were estimated in non-overlapping windows of 20 and 100 kb with the following filters: minimum base quality of 20, minimum coverage of 30, minor allele count of 3 and 3 SNPs/window (F_{ST} was also estimated by SNP for the identification of outliers, described below). Twenty kb was the smallest window we could use without losing a large number of windows to the filter requiring 3 SNPs/window (e.g., 4002 windows lost at 10 kb; only 432 lost at 20 kb [<1%]) and the 100 kb window size is comparable to that used in other songbird studies (Ellegren et al. 2012). In the candidate gene analysis described below we also used windows of 1000 kb to allow comparisons with results from Ruegg et al. (2014). We did not use this window size in other analyses because it summarized over a very large number of SNPs in our study (mean number of SNPs = 2480 in 1000 kb windows; 252 in 100 kb windows; 51 in 20 kb windows).

Note that there is an additional source of sampling error when estimating F_{ST} from pooled samples (vs. data from individuals): error introduced when reads are drawn from the pool of individuals sampled from the population. There is currently no established method to correct for this source of error, which results in estimates of F_{ST} that are biased upward. Because this bias should have a roughly similar effect across the genome in a single analysis, relative comparisons among SNPs in each comparison can be made but direct comparisons with other studies cannot. The filters we applied to SNPs before they were included in our analysis (especially the high minimum coverage cutoff) helped reduce the impact of this sampling error (and sequencing errors, Futschik and Schlötterer 2010; Kofler et al. 2011a). Note that several additional factors, including differences in sampling design, can also affect estimates of F_{ST} and can preclude comparisons between studies. For example, Ruegg et al. (2014) combined data from

geographically distant populations prior to estimating F_{ST}. Genetic differentiation between these populations may have artificially increased within population differentiation, reducing estimates of F_{ST}. The populations included in our study were more proximate and should not be as affected by this phenomenon.

4.2.4 Identification of outliers

We used F_{ST} estimated at each SNP with a Hidden Markov Model (HMM) to identify outlier windows. This model has been used by several recent papers to quantify genomic differentiation (Hofer et al. 2012; Hübner et al. 2013; Soria-Carrasco et al. 2014). The goal is not to categorize the differentiation at specific SNPs but rather to identify chromosome regions that are highly differentiated. Briefly, each SNP is assigned to one of three hidden states – high, moderate or low differentiation. The state of each SNP is dependent on that of the previous SNP on the chromosome. Initial transition (probability of switching between states) and emission (probability of high, moderate and low states) matrices are provided to the model. The model uses these matrices to predict which state was most likely to have generated the observed value of F_{ST} at each SNP. The main benefit of this model is that it incorporates the non-independence of proximate SNPs. We used the 'HiddenMarkov' package in R 3.0.3 (R Development Core Team 2014) and logit transformed F_{ST} to approximate a Gaussian distribution. We generated our initial emission matrix using the 'compdelta' function and used a transition matrix with a higher probability of staying in the current state (vs. switching states) and zero probability of moving between high and low states. We used the Baum and Welch algorithm to estimate the mean logit F_{ST} for each state, transition and emission matrices and a local decoding approach to assign each SNP to a state based on its highest conditional expectation.

We summarized results from the HMM into non-overlapping windows of 20 and 100 kb. Windows with a substantially higher proportion of SNPs assigned to the high differentiation state than the rest of the genome were considered outliers. We used a resampling test to identify these outliers: for each window we randomly sampled the same number of SNPs from the rest of the autosomal genome and calculated the proportion of "high" SNPs in the sample. We used 1000 resamples to simulate a null distribution and calculated *P*-values as the fraction of resampled proportions that were greater than the observed value. To improve accuracy in the tail of our distribution, we subjected all windows with $P < 0.1$ to an additional 100,000 resamples (Andrew et al. 2012; Renaut et al. 2013). We controlled for multiple testing using the p.adjust function in R (p.adjust = 0.01, using the "fdr" method, which controls for false discovery rates when estimating significance). We note that the goal of this approach is to focus subsequent analyses on areas of the genome that have the most evidence for being under selection; our goal is not to determine precise *P*-values of the data under a null model of neutrality, as the latter would require methods to incorporate the demographic history of these populations and recombination structure of the genome. In addition, we have tried various summary statistics (e.g., proportion of SNPs with F_{ST} values in the top 3%, proportion of significant SNPs based on Fisher's exact tests [Appendix C.4]) for this analysis and the results remained qualitatively the same. We did not include the Z chromosome in this analysis as this chromosome is inherited differently than autosomes and most of the Z chromosome was strongly differentiated (see Results).

4.2.5 Testing associations between patterns of genomic variation and features of the genome

To clarify patterns of within and between population genomic variation, we compared values of

d_{XY}, π and Tajima's *D* inside and outside outlier windows using permutation tests implemented in the 'coin' package of R with 10,000 resamples. We used the same method to compare values of F_{ST} on autosomal vs. sex chromosomes and inside vs. outside centromeric regions. Methods for identifying centromeric regions are described in Appendix C.1; note that information on the location of centromeres was only available for nine chromosomes and thus comparisons of F_{ST} were limited to these chromosomes.

4.2.6 Candidate genes and enriched functional categories

In addition to elevated F_{ST}, strong divergent selection is also expected to alter the site frequency spectrum of segregating allelic variation (Tajima 1989). Of particular interest to this study, strong directional selection tends to increase the number of low-frequency alleles compared to high-frequency alleles and lead to negative values of the summary statistic Tajima's *D* (Tajima 1989). To target regions of the genome experiencing strong divergent selection, we limited the remainder of these analyses to windows that were outliers in the F_{ST} analysis and within the bottom 5% threshold for Tajima's *D* in each subspecies group.

We searched for genes from two lists of candidate genes. The first list was compiled by Ruegg et al. (2014) based on a literature review of genes linked to migration. We generated the second list using the search terms 'cryptochromes', 'circadian rhythm' and 'biological rhythm' in Ensembl 75 and limiting our results to genes in the zebra finch genome. Regarding our choice of search terms, taxa that form migratory divides often differ in their migratory timing, distance and direction (Irwin and Irwin 2005). Circadian rhythms are involved in determining migratory timing and distance (Gwinner 1996) and biological compasses likely help determine migratory direction (Berthold 1991). The magnetic compass has received a good deal of support (Cochran

et al. 2005) and may involve the use of cryptochromes, light-dependent receptors located in the eye (Liedvogel et al. 2011). Following Schielzeth et al. (2012), in this second list we also included genes that were identified in this initial search and functionally related to them to account for the fact that our functional knowledge of migration is still quite general. In one example, two dopamine receptors (DRD2 and DRD3) were identified in our initial search. Three additional genes coding dopamine receptors are located in the zebra finch annotation. One of these genes, DRD4, is involved in exploratory behavior in many taxa (Fidler et al. 2007), suggesting that the inclusion of functionally related genes may have improved our analysis.

We searched for candidate genes in the outlier windows described above and limited our results to genes containing at least one SNP assigned to the high differentiation state based on the HMM analysis, as it is likely that some of the genes in these divergence peaks are not the targets of selection. We used binomial tests to compare the proportion of genes from each of our two lists of candidates to what would be expected if candidate genes were randomly distributed across the regions of the genome we have in our assembly. It is possible that genes are not randomly distributed across the genome. Accordingly, we also used a Chi-square contingency test to compare the number of candidates inside and outside outliers to the total number of genes in these areas.

We used GOwinda (Kofler and Schlötterer 2012) to determine if SNPs within divergence peaks showed enrichment for any gene ontology (GO) categories. This program controls for gene length bias and requires two lists of SNPs as input: a list of all SNPs and those that are candidates for selection. We considered SNPs candidates if they were assigned to the high differentiation state based on the HMM model described above. We included SNPs that were 2000 bp upstream and downstream from genes to account for the fact that some SNPs may be

enriched in regulatory regions, and we assumed SNPs within a gene were in complete linkage disequilibrium. We used a false discovery rate of 0.05 and required each GO category to have at least 3 genes associated with it.

4.3 Results

4.3.1 Draft reference genome

The number of raw and trimmed reads for both fragment and mate pair libraries are summarized in Appendix C.2. Coverage of the fragment and mate pair libraries (2 and 4 kb) was approximately 30x, 11x and 10x, respectively. The final Swainson's thrush assembly is 1,233,099,904 bp in length, with 304,366,159 N's and 13,482 genes. This assembly includes sequence for all of the assembled zebra finch chromosomes and linkage groups. Summary statistics for the assembly are provided in Table 4.1. To obtain statistics similar to those from *denovo* assemblies, we broke the genome at strings of 10 Ns (roughly equivalent to contigs [roughly contiguous sequences]), 2000 and 4000 Ns (roughly equivalent to scaffolds [contigs connected by strings of Ns] from the 2 and 4 kb libraries).

The length of our assembly is similar to other avian genomes (all between 1 and 1.2 Gb, Ellegren 2013). We have fewer genes in our annotation (e.g., 18,375 genes in the collared flycatcher assembly, Ellegren et al. 2012), likely reflecting the fact that there are still large gaps in our assembly. Faircloth et al. (2012) identified ultra-conserved elements using whole genome alignments for the chicken, anole and zebra finch. To validate our genome, we downloaded the sequences for these elements and used NCBI's 'blastn' to see how many were in our assembly. 5081 of the 5561 elements (91.4%) are in our assembly. These authors included a more limited set of elements with greater coverage. Of these 2560 elements, our assembly has 2544 (99.4%).

4.3.2 Relative differentiation and outliers

The number of raw, trimmed and mapped re-sequencing reads for each pool is summarized in Appendix C.2. Between 67 and 71% of the trimmed reads from each library mapped to the reference, resulting in a median coverage for the coastal and inland subspecies groups of approximately 17x and 20x, respectively (Appendix C.5). Following filtering, 2,571,420 SNPs between coastal and inland subspecies groups of Swainson's thrushes were identified. Many genomic studies of songbirds rely on the zebra finch genome to map their reads and call SNPs. To quantify the increase in resolution gained by assembling a draft reference genome for the Swainson's thrush, we used the same pipeline to align re-sequencing reads to the zebra finch genome and call SNPs. There was a 16% increase in the number of reads mapped to the thrush genome (67-71% with our draft genome vs. 51-54% with zebra finch.), resulting in a ~2 fold increase in the number of SNPs called (2,571,420 vs.1,264,664 SNPs).

We observed a great deal of structure in relative differentiation across the genome. Figure 4.1 shows windowed estimates of F_{ST} across the entire genome and Figure 4.2 shows the distribution of these estimates, with values ranging from 0 to 1 (mean 0.10, median 0.08). We used a HMM to categorize SNPs into different states of differentiation and a resampling test to identify outlier windows (i.e., those with a higher proportion of SNPs assigned to the high differentiation state). Summarizing data by 100 kb windows, 2273/9494 windows were considered outliers (gray lines, Figure 4.1) and had average and median F_{ST} values of 0.18 and 0.16, respectively (vs. 0.079 and 0.075 in non-outlier windows). These outlier windows combined into 636 peaks (i.e., continuous strings of significant windows), with a mean and median size of 357,400 and 100,000 bp, respectively. Summarizing data by 20 kb windows,

6362/46652 windows were identified as outliers (gray lines, Appendix C.6) and had average and median F_{ST} values of 0.23 and 0.16, respectively (vs. 0.086 and 0.079 in non-outlier windows). These windows combined into 3262 peaks, with a mean and median size of 39,010 and 20,000 bp, respectively. Almost all of the chromosomes had at least one outlier peak and most of the larger ones (i.e., macrochromosomes) had more than one (Figure 4.1, Apprendix C6).

4.3.3 Associations between patterns of genomic variation and features of the genome

Estimates of d_{XY}, π and Tajima's D are shown in Figures 4.3 and Appendix C.7, along with estimates of F_{ST}. These parameters are variable across the genome and show a strong relationship with F_{ST}. Comparing values within and outside outlier windows, π and Tajima's D are significantly lower in outlier windows (π: 0.0039 [coastal] and 0.0052 [inland] in outliers vs. 0.0094 [coastal] and 0.010 [inland] in non-outliers; Tajima's D: -0.76 [coastal] and -0.78 [inland] in outliers vs. -0.14 [coastal] and -0.33 [inland] in non-outliers; all permutation tests $P < 0.001$). Intriguingly, not only were values of d_{XY} not higher in outlier windows, but they were actually lower (d_{XY}: 0.0059 in outliers vs. 0.0098 in non-outliers, permutation test $P < 0.001$). This pattern was even more apparent in outlier windows identified using both the HMM and Tajima's D (see below; 0.0034 in outliers vs. 0.0092 in non-outliers, permutation test $P < 0.001$).

Figures 4.3 and Appendix C.7 also include F_{ST} estimated between populations of the same subspecies groups, which shows far less variability across the genome compared to F_{ST} estimated between the groups (standard deviation of $F_{ST} = 0.06$ between coastal and inland, vs. 0.03 and 0.02 for F_{ST} estimated between coastal and between inland populations, respectively). Note that there are regions where F_{ST} values estimated within groups exceed those between; this is likely related to the fact that the sample size is halved in the within-group comparisons,

increasing the bias in F_{ST} resulting from sampling error (see 'Methods'). Accordingly, patterns across the genome rather than actual values of F_{ST} should be compared between vs. within the groups.

Between groups, estimates of F_{ST} on the Z chromosome were significantly higher than those on the autosomal chromosomes (F_{ST} = 0.17 on Z chromosome and 0.09 on autosomal chromosomes, permutation test $P < 0.001$; Figure 4.1, 4.2a) and there were significantly more fixed SNPs on the Z chromosome (Chi-square contingency test, $\chi^2 = 1.4x10^4$, $P < 0.001$; Figure 4.2a). Estimates of F_{ST} in centromeric regions (excluding the Z chromosome) were also significantly higher than those in non-centromeric regions (F_{ST} = 0.11 in centromeric regions and 0.10 in non-centromeric regions; permutation test $P < 0.001$; Figure 4.2b), although the difference in mean F_{ST} was quite small. Similar to the comparison between autosomes and the Z chromosome, there were significantly more fixed SNPs in centromeric regions (Chi-square contingency test, $\chi^2 = 742.4$, $P < 0.001$; Figure 4.2b). To evaluate the relationship between F_{ST} and centromeres further we used a Chi-square contingency test to compare the number of windows in centromeric regions that were outliers (212) and non-outliers (459). There was a significantly higher proportion of outlier windows in centromeric regions than would be expected by chance ($\chi^2 = 68.4$, $P < 0.001$). Running counter to this general association, the centromeric regions on chromosomes 4 and 5 did not include any outlier windows and the centromeric region on chromosomes 3 and 7 included only one and two outlier windows, respectively (Figure 4.1).

4.3.4 Candidate genes and enriched functional categories

To focus on regions of the genome most likely to be experiencing strong directional selection we limited subsequent analysis to outlier windows from the HMM described above that were also in the bottom 5% of Tajima's *D* estimates in both subspecies groups (red lines Figures 4.1 and Appendix C.6). This 5% cut-off corresponded with values of -1.33 and -1.36 for Tajima's *D* in 100 and 20 kb window analyses, respectively. Of the 9494 windows of 100 kb, 563 were F_{ST} outliers and in the bottom 5% of Tajima's *D* estimates. These windows grouped into 179 peaks (Figure 4.1), which included 983 genes, 697 of which had at least one SNP assigned to the high differentiation state in the HMM. Of the 46,652 windows of 20 kb, 1892 were outliers following both criteria. These windows grouped into 1042 peaks (Appendix C.6), which included 804 genes, 651 of which included at least one highly differentiated SNP.

Of the 25 genes listed by Ruegg et al. (2014) as candidates for migratory behaviour, 21 were in our annotation (vs. 5 in Ruegg et al. [2014]); two of these 21 genes were located on the Z chromosome and thus excluded from subsequent analyses [Table 4.2]). Four of the remaining 19 candidates were in 100 kb outlier windows, and four were in 20 kb outlier windows (Table 4.2). Two of these genes were common to both windowed analyses. The six genes (2 in both windowed analyses and 4 in just one or the other) are spread throughout the genome; two were on the same chromosome but located in different outlier peaks (Table 4.2). For both window sizes, the number of candidates in outlier windows was greater than would be expected if genes were distributed randomly (binomial test: expected number = 1.04 and 0.97, P = 0.02 and 0.01 for 100 and 20 kb analyses) or randomly among the locations of genes in the genome (Chi-square contingency test: χ^2 = 3.47 and 5.20, P = 0.03 and 0.01 for 100 and 20 kb analyses). To compare our results with those of Ruegg et al. (2014) we also looked for candidates in outlier

windows of 1000 kb (Appendix C.9). Only one of the 19 candidates was within these windows (Table 4.2). This gene, CPNE4, was also the only gene identified as being in an outlier peak by Ruegg et al. (2014). Note that this gene was also within our 100 kb outlier windows but not within our 20 kb outlier windows (Table 4.2). In contrast to the significant enrichment of candidate genes in 20 kb and 100 kb outlier windows, the number of candidate genes in 1000 kb outlier windows was not more than would be expected by chance (binomial test: $P = 0.59$; Chi-square contingency test: $\chi^2 = 0.03$, $P = 0.43$).

In a broader search for candidate genes, we identified 100 candidates using search terms related to migratory timing, distance and direction and including functionally related genes; 73 were in our annotation and one was on the Z chromosome and thus excluded from subsequent analyses (Appendix C.8). Nine of these 72 genes were in 100 kb outlier windows, 11 were in 20 kb outlier windows and 8 of these genes were in both outlier window sizes (Appendix C.8). Of these 12 genes (8 in both windowed analyses, 1 and 3 in 100 and 20 kb analyses), 5 were located on separate chromosomes and seven were located on the same chromosome as at least one other gene (Appendix C.8). Of these seven genes, the three on chromosome 10 were in the same peak but the rest were in different peaks. The number of candidates in both window sizes was more than would be expected if candidate genes were distributed randomly in the genome (binomial test: expected number = 3.93 and 3.66, $P = 0.02$ and 0.001 for 100 and 20 kb analyses) or randomly among locations of genes in the genome (Chi-square contingency test: $\chi^2 = 2.15$ and 5.20, $P = 0.07$ and 0.001 for 100 and 20 kb analyses), although significance was not met for the latter analysis with 100 kb windows. For the analysis of 1000 kb windows, 8/73 candidates were found in these windows; this number was more than expected by chance (binomial test, $P = 0.01$).

For autosomes that had at least one candidate gene in an outlier peak, Figures 4.3a and Appendix C.7 show the location of all genes in the annotation (orange) and candidate genes that are within outlier peaks (blue). Many of the candidates that were in outlier peaks occur at the center of potential selective sweeps, where estimates of F_{ST} are near their highest and estimates of both π and Tajima's D are near their lowest. Figure 4.4 narrows in on DRD3, a candidate found in an outlier peak on chromosome 1 in each windowed analysis. DRD3 encodes a dopamine receptor; these receptors are associated with exploratory behaviour in many taxa (Fidler et al. 2007). This gene had three non-synonymous SNPs. Two of these SNPs were fixed or nearly fixed for different alleles in the subspecies groups, and the third has little variability. Table 4.2 and Appendix C.8 include columns summarizing the proportion of non-synonymous SNPs found in each candidate gene located in an outlier window. Eleven of 14 of the candidates identified in outlier windows included at least one non-synonymous SNP.

So far we have focused on candidate genes found in outlier peaks. We also compared F_{ST} values of SNPs within candidate genes vs. F_{ST} of SNPs within genes that were not in our lists using a resampling test (1000 resamples). Candidates from the list generated by Ruegg et al. (2014) had significantly higher values of F_{ST} (0.17 vs. 0.11, $P = 0.015$); this was not the case for the longer candidate list generated in the present study (0.10 vs. 0.11, $P = 0.78$).

SNPs within outlier peaks that were classified as highly differentiated by the HMM analysis were significantly enriched for two GO categories, spindle assembly (biological process) and kinetochore binding (molecular function; Appendix C.10). This indicates that outlier peaks tend to have more genes that are associated with these two functions (both of which are involved in cell division) than expected if these gene functions were spread randomly among all genes in the genome.

4.4 Discussion

We documented substantial genomic heterogeneity in sequence divergence between two subspecies groups of the Swainson's thrush, with highly differentiated SNPs located throughout the genome and in clusters of varying sizes. This finding is in accordance with Ruegg et al. (2014), who used RAD sequencing data from populations more distant from the hybrid zone, and contrasts with studies of other migratory divides with extensive interbreeding, which have found very low levels of genetic differentiation (e.g., willow warblers, Bensch et al. 2009; Lundberg et al. 2013; European blackcaps, Pérez-Tris et al. 2004; Mettler et al. 2013). Combined with over a decade of work on reproductive isolation between these subspecies groups (Ruegg and Smith 2002; Ruegg et al. 2006a,b; Ruegg 2008; Delmore et al. 2012; Delmore and Irwin 2014), these results strengthen the Swainson's thrush as an important system for studying adaptation and speciation in birds, especially in relation to seasonal migration.

Inherent features of the genome likely contribute to the heterogeneous pattern we reported. Estimates of F_{ST} were modestly higher in centromeric regions, more fixed SNPs were located in these regions and they included more F_{ST} outlier windows. We note that the association between centromeres and relative differentiation may be even stronger than we reported, as microchromosomes are generally acro- or telocentric and many divergence peaks were located at the ends of these chromosomes (Figure 4.1). Estimates of F_{ST} were also much higher on the Z chromosome and this chromosome included more fixed differences than autosomal chromosomes. Lower rates of recombination can lead to reductions in within population polymorphism through linked selection (Hellmann et al. 2003; Kulathinal et al. 2008)

and likely accounts at least partially for greater differentiation near centromeres and on the Z chromosomes.

The striking increase in relative differentiation documented on the Z chromosome is likely generated by additional factors beyond just reduced recombination. For instance, if the fitness of hybrids is influenced by interactions between recessive genes, genes on the sex chromosomes, which occur in a hemizygous state in the heterogametic sex, will experience lower rates of gene flow between hybridizing populations compared to those on the autosomes (Mank et al. 2007; Presgraves 2008). Speciation genes may also occur at a higher density on the Z chromosome (Presgraves 2008), especially if they are involved in sexual selection (Albert and Otto 2005). The latter suggestion is specific to ZW-systems of sex determination (e.g., birds) and follows from the fact that Z-linked traits will be passed directly from father to son in these systems. Accordingly, sexually antagonistic alleles that may reduce the fitness of females (e.g., those influencing large ornamental traits) do not have to pass through daughters before they appear in another male (Albert and Otto 2005). Several other studies have highlighted the role the Z chromosome may play in speciation of birds, including genome scans (Ellegren et al. 2012), comparisons of gene flow between the autosomes and the Z chromosome (Carling and Brumfield 2009; Storchová et al. 2010), and studies identifying specific genes involved in speciation (Saetre et al. 2003).

Associations with centromeres and the Z chromosome cannot explain all of the patterns we observed; some of the centromeric regions did not have outlier windows and several autosomes had more than one set of outliers that were not located near centromeres. The relationships we documented between relative differentiation (F_{ST}) and within population diversity (π and Tajima's D) strongly suggest that variability in the strength of selection could be

helping generate these patterns. Reductions in within population diversity and the number of high frequency polymorphisms (i.e., negative values of Tajima's *D*) are indicative of selective sweeps (Maynard Smith and Haigh 1974; Tajima 1989) and these parameters were lower in islands of relative differentiation. Reductions in these parameters were observed in both inland and coastal populations, suggesting that sweeps occurred in both subspecies groups. In addition, in contrast to the expectation that absolute differentiation would be elevated in islands of relative differentiation, estimates of d_{XY} were surprisingly lower in islands; this suggests that heterogeneity in relative differentiation is not primarily driven by gene flow, which should result in both lower absolute (d_{XY}) and relative differentiation (F_{ST}) in the regions experiencing high gene flow. Combined, these findings argue for an important role of selective sweeps in generating islands of relative differentiation, rather than variation in the role of different genomic regions in causing reproductive isolation and thereby reducing gene flow (Cruickshank and Hahn 2014). The divergence-with-gene-flow model has received a great deal of attention in the literature (Nosil et al. 2009; Ellegren et al. 2012; Feder et al. 2012; Via 2012); similar applications of absolute differentiation may help highlight the role selective sweeps alone could play in generating patterns of heterogeneity.

Previous suggestions that differences in seasonal migration play a role in reducing gene flow between Swainson's thrushes have been supported by tracking data showing that hybrids take intermediate routes over several geographic barriers (Delmore and Irwin 2014), and patterns of molecular variation in the hybrid zone also suggest strong selection against hybrids (Ruegg 2008; Delmore and Irwin 2014). Results from the present study support these findings, with a higher proportion of genes linked to migratory behaviour found in outlier windows. Values of F_{ST} were higher within candidate genes listed by Ruegg (2014), and many genes found within

outlier windows were located near the center of potential selective sweeps (i.e., areas where estimates of F_{ST} were near their highest and within population diversity their lowest) with non-synonymous SNPs. Even the SNPs that are synonymous could be subject to selection, since they can cause variation in translation efficiency and the stability of mRNA (Shields et al. 1988; Cuevas et al. 2012). SNPs may also occur within regulatory regions of these genes. Several genes involved with the circadian clock, which influences the timing and distance of migration (Gwinner 1996; Bartell and Gwinner 2005) were found in these outlier windows. For example, the CLOCK gene encodes transcriptional factors that help activate period and cryptochrome genes, including PER2 and CRY1. Once proteins from the latter genes reach critical concentrations, they de-activate CLOCK, generating daily oscillations. CSNK1E helps regulate this loop by phosphorylating period genes (Kloss et al. 1998; Wager-Smith and Kay 2000). The list of candidates generated by the present study may have been too broad (as evidenced by the lack of an increase in F_{ST} within these candidates compared to other genes). Nevertheless, genes from this list were enriched in outlier peaks and are promising targets for future analyses on the genetic basis of migration, as are genes in outlier peaks associated with enriched GO categories.

The association we documented between genes linked to migration and outlier windows expands on results presented by Ruegg et al. (2014), who found one candidate gene in their outlier windows (i.e., no significant association). The increased resolution obtained in the present study, with the use of WGS sequencing and assembly of a draft reference genome, likely permitted us to find this association. Our study included an order of magnitude more SNPs (2,571,420 vs. 154,123 SNPs), providing more statistical power and greater coverage of genes (13,482 genes were in our annotation vs. 5,183 in Ruegg et al. [2014]; 21/25 potential candidate genes listed by Ruegg et al. [2014] were in our annotation vs. 5 in Ruegg et al. [2014]). The

increase in SNPs also allowed us to use window sizes that were 10-50 times smaller than Ruegg et al. (2014), increasing the spatial resolution of differentiation. This point is illustrated well by the fact that CPNE4 was identified as being in an outlier window by Ruegg et al. (2014) and in our analyses using large windows (1000 and 100 kb) but was lost as an outlier in analyses using 20 kb windows, suggesting that a different gene in the vicinity of CPNE4 may actually be the target of selection. These methodological differences led to different conclusions in the two studies, with Ruegg et al. (2014) concluding no association between migratory genes and islands of differentiation, whereas we have found an association. Additional analyses will be needed to establish a direct link between these genes and migratory traits. For example, several hybrid populations have been described between coastal and inland Swainson's thrushes in the Coast Mountains of British Columbia (Ruegg 2008; Delmore and Irwin 2014) and would be well suited to admixture mapping and analyses of geographic and genomic clines (Barton and Hewitt 1985; Gompert et al. 2012).

Using next generation sequencing techniques and focusing on a well characterized avian system, we have provided important insight into the genomic architecture of divergent selection and reproductive isolation. Our results support a role for reductions in within population diversity accompanying selective sweeps in generating genomic heterogeneity in relative differentiation. Cruickshank and Hahn (2014) emphasized the role of selective sweeps following geographic isolation in generating genomic heterogeneity. In their re-analysis of genomic datasets, they also observed that in some cases d_{XY} was actually lower in areas of high F_{ST}. They suggested that recurrent selection in the ancestral population could be responsible for this pattern, as it would reduce d_{XY} prior to divergence. This explanation could also account for the patterns we observed, with significantly lower d_{XY} in islands of relative differentiation.

Nevertheless, we propose a variant of this basic idea to explain reductions in d_{XY}, in which populations come into contact at various times during their divergence from the ancestral population. This scenario is likely for many systems, especially those in temperate regions where glacial cycles have likely led to periods of both allopatry and secondary contact (Hewitt 2000). Following geographic isolation, one or both populations could acquire globally adaptive mutations. Following secondary contact, these alleles could sweep from one population to the other, reducing d_{XY} to zero. Local adaptation and selective sweeps following the re-establishment of geographic isolation could then cause F_{ST} to increase and within population diversity to decrease. This process could take place several times throughout the geographic history of a system, accounting for the reductions in d_{XY} we documented. This model remains speculative but is interesting to consider, as it matches the geographic history of many populations better and makes the opposite prediction to the divergence-with-gene flow model, as we are proposing that areas of high F_{ST} have actually experienced gene flow more recently than the rest of the genome.

Table 4.1 Summary statistics for draft reference genome broken at strings of 10, 2000 and 4000 Ns. Breaks at strings of 10 roughly approximate statistics for contigs from *denovo* assemblies; strings of 2000 and 4000 scaffolds from *denovo* assemblies

Breaks (# Ns)	Length (bp)	# Sequences	N50	# Ns	GC %
10	931,262,738	1,733,030	1,050	2,528,993	40.6
2000	1,193,845,719	10,399	621,829	265,111,974	31.7
4000	1,215,699,857	2,238	5,201,219	286,966,112	31.1

Table 4.2 List of candidate genes associated with migratory behaviour generated by Ruegg et al. (2014). Information on the gene's location in our annotation (Chrom, - if not present) and whether it was present in Ruegg et al. (2014) is provided, along with whether the gene was found in outlier windows of 1000, 100 and 20 kb. The proportion of non-synonymous (NS) SNPs in each gene found in an outlier window is also provided.

Gene name	Chrom	Ruegg 2014	1000 kb	100 kb	20 kb	NS SNPs
bmal1/ARNTL	1A	N	N	N	N	
CRY1	1A	N	N	N	**Y**	0/2
CRY2	5	N	N	N	N	
NA/PER3	21	Y	N	N	N	
TTR/B5G258_TAEGU	2	N	N	N	N	
YPEL1**	-	N	n/a	n/a	n/a	
GRP94/ HSB90B1	-	N	n/a	n/a	n/a	
HSP90AA1	3	N	N	N	N	
Bip/HSPA8	24	N	N	N	N	
HSPA5	17	N	N	N	N	
GLUT1/SLC2A1	21	N	N	**Y**	N	0/5
PARL	9	N	N	N	N	
HRSP12	2	N	N	N	N	
SLC1A3	Z	N	n/a	n/a	n/a	
NIMA/NEK2	-	N	n/a	n/a	n/a	
CPNE4	2	Y	**Y**	**Y**	N	1/6
CREB1	-	Y	n/a	n/a	n/a	
NPAS2	1	Y	N	N	N	
DRD4	5	N	N	N	N	
ADCYAP1	2	Y	N	N	N	
CLOCK	4	N	N	**Y**	**Y**	1/3
CSNK1E	1A	N	N	**Y**	**Y**	2/3
AANAT	18	N	N	N	N	
PER2	9	N	N	N	**Y**	3/8
NFIL3/E4BP4	Z	N	n/a	n/a	n/a	

Figure 4.1 Mean F_{ST} in 100 kb, non-overlapping windows between populations of coastal and inland subspecies groups of Swainson's thrushes. Gray lines at the top of each panel show F_{ST} outlier windows identified using a HMM; red lines identify F_{ST} outlier windows that were also in the bottom 5% of Tajima's *D* values within both subspecies groups. Blue lines at the bottom of panels show the inferred locations of centromeric regions. F_{ST} was summarized for the Z chromosome but this chromosome was not included in outlier analyses.

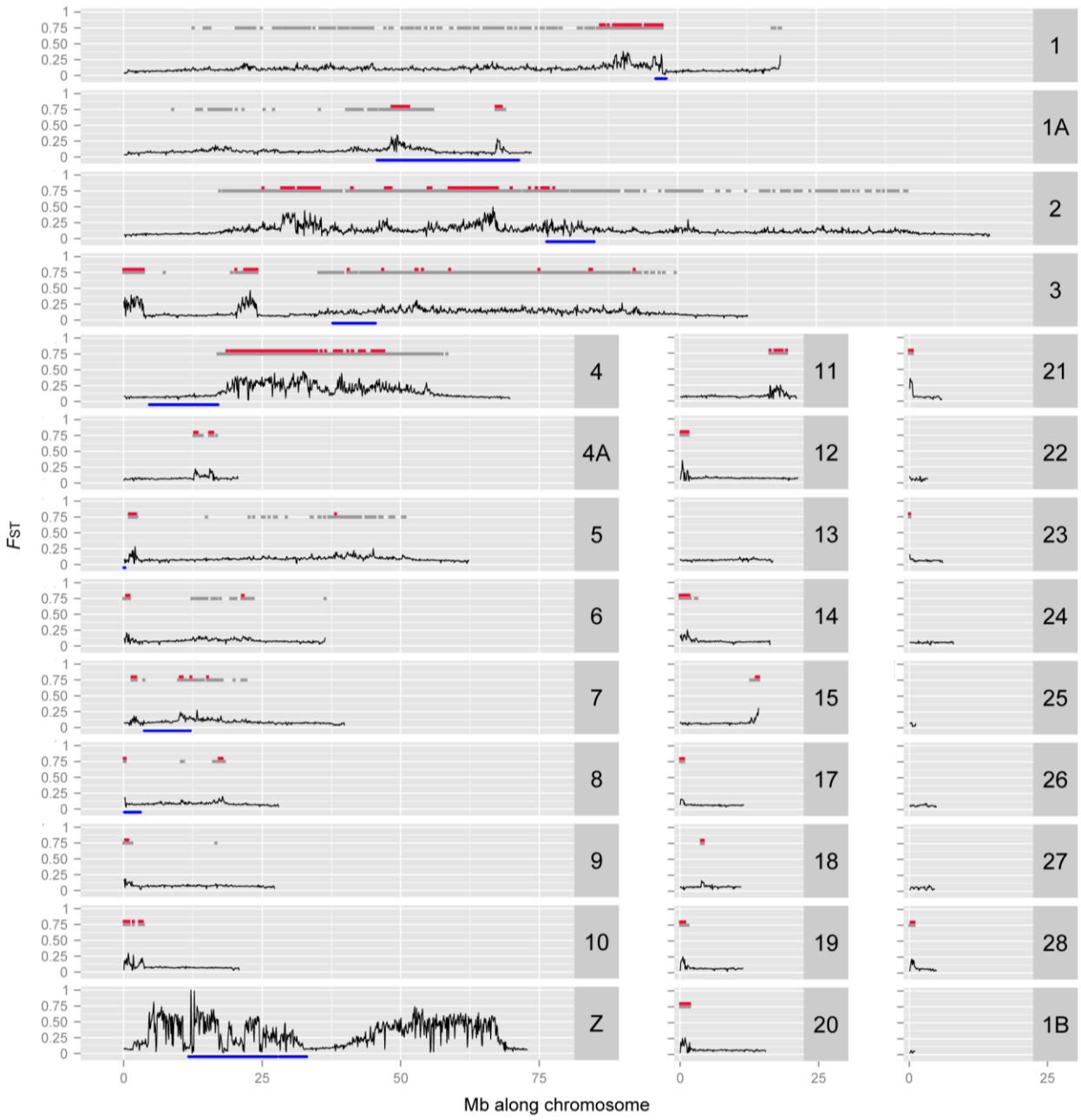

FST
1
0.75
0.50
0.25
0
1A
2
3
4
4A
5
6
7
8
9
10
Z
11
12
13
14
15
17
18
19
20
21
22
23
24
25
26
27
28
1B
0
25
50
75
Mb along chromosome

Figure 4.2 Histograms showing the proportion of SNPs with F_{ST} values in thirty equally sized bins between estimates of 0 and 1. A comparison of F_{ST} estimated on the autosomes (gray) and Z chromosome (black) is shown in panel a and a comparison between loci located inside (black) and outside (gray) centromeric regions is shown in panel b.

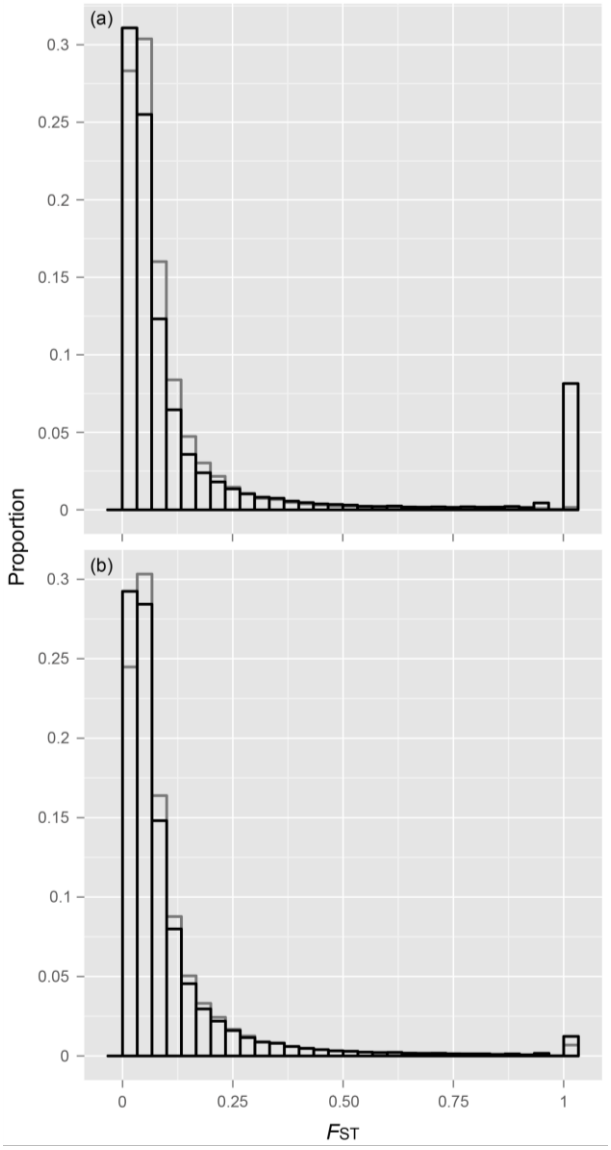


Figure 4.3 Population genomic estimates for chromosome 1. Panel (a) is an enlargement of Figure 4.1 for this chromosome, including mean estimates of F_{ST} between the subspecies groups, outliers identified using F_{ST} (gray), outliers identified using F_{ST} and Tajima's D (red), all of the genes on this chromosome (orange) and candidate genes found in non-overlapping outlier windows of 100 kb (blue, KIAA2018 and DRD3). The remaining panels show d_{XY} (b), F_{ST} estimated between populations of the same subspecies groups (c, estimates between coastal populations in blue, between inland populations in yellow), Tajima's D (d, estimates for inland populations in purple, for coastal populations in green) and nucleotide diversity (e, estimates coloured as in panel d).

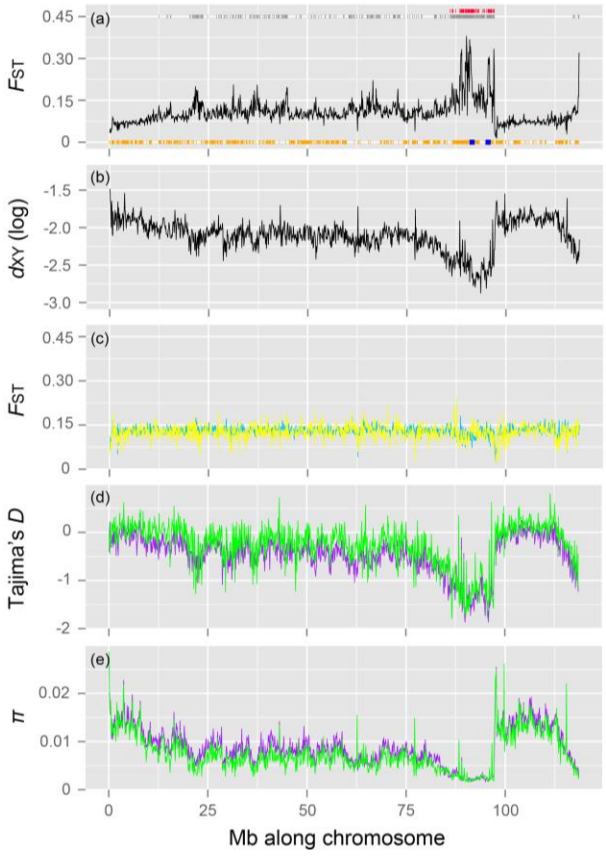
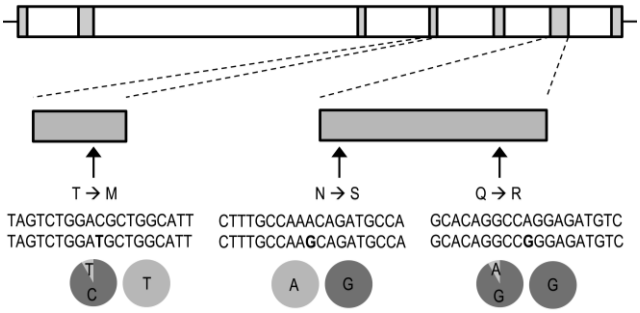


Figure 4.4 Expansion of DRD3, a candidate gene found in an outlier window on chromosome 1 and shown in Figure 4.3. The gene is 11,519 bp in length. Dark gray indicates exonic regions of the gene, arrows show the position of non-synonymous SNPs (with amino acid changes as indicated), and pie charts illustrate the proportion of each base in each population (coastal on left, inland on right).



Chapter 5: Phenotypic divergence during speciation is inversely associated with differences in seasonal migration

5.1 Introduction

Seasonal migration has been proposed as an important driver of reproductive isolation and speciation. This suggestion derives from cases of 'migratory divides,' in which closely related but somewhat differentiated populations breed adjacent to one another but use different migratory routes (Veen et al. 2007; Bensch et al. 2009; Rolshausen et al. 2009). Differences in the migratory behaviour of these populations could serve as both pre- and postmating isolating barriers (reviewed by Irwin and Irwin 2005). For instance, populations at divides could arrive on their breeding grounds at different times and mate assortatively based on arrival time (Irwin and Irwin 2005; Webster and Marra 2005). Migratory routes are also largely genetically determined in many groups and can involve navigation around unsuitable areas (e.g., deserts and mountain ranges). Accordingly, hybrids in divides could use intermediate or mixed routes that are inferior to those of parental forms (Helbig 1991; Bearhop et al. 2005; Irwin and Irwin 2005; Veen et al. 2007; Bensch et al. 2009; Rolshausen et al. 2009; Møller et al. 2011; Rohwer and Irwin 2011; Delmore and Irwin 2014).

Migratory divides are found in many taxonomic groups (O'Corry-Crowe et al. 1997; Gagnaire et al. 2009; Altizer and Davis 2010) and there is growing interest in the role that they and other behavioural differences could play in speciation (Coyne and Orr 2004; Price 2008). A few empirical studies focusing on single species have accumulated support for migration's role in speciation. For example, Bearhop et al. (2005) used stable isotopes to identify the wintering

grounds of European blackcap populations that form a divide in central Europe. These authors documented differences in arrival time between the populations and assortative mating based on location of their wintering grounds. In another study, Helbig (1991) bred first generation hybrids between populations of blackcaps and assayed their orientation using Emlen funnels. Hybrids oriented in directions intermediate to parental forms that would have brought them over the Alps, the Mediterranean Sea and the Sahara Desert. More recently, Delmore and Irwin (2014) used light-level geolocators to track hybrids between two subspecies groups of Swainson's thrushes that form a divide in western North America; hybrids took intermediate routes over mountainous and arid regions in southwestern North America. Finally, Bensch et al. (1999, 2009) documented strong associations between migratory behaviour (quantified using stable isotopes and band recoveries) and clines in various morphological and genetic traits across a narrow hybrid zone between two subspecies of willow warbler in Scandinavia.

A few studies have begun to examine the role migration plays in speciation on a broader scale. For example, Møller et al. (2011) used data on band recoveries to identify migratory divides between passerines in Europe. Over one quarter of all species they identified formed divides and many showed evidence of local adaptation, likely reducing gene flow at divides. Rohwer and Irwin (2011) identified migratory divides between birds in North America, and found that twelve pairs of closely related taxa had well-studied contact zones and eight formed migratory divides with substantial reproductive isolation (a lack of hybrids and/or narrow hybrid zone). Likewise, north Asian migratory behaviour was reviewed by Irwin and Irwin (2005), who found that seven species are known to form migratory divides between western and eastern subspecies that migrate to south Asia via routes going west or east of the Tibetan Plateau, a high-elevation desert. These studies established the plausibility that migratory divides could contribute

to reproductive isolation on a broad scale but did not include a rigorous statistical analysis or a quantitative measure of reproductive isolation.

Quantitative measures of reproductive isolation are often collected using mate choice trials and measurements of the viability and fertility of hybrids produced in lab crosses (Coyne and Orr 1997; Price and Bouvier 2002; Lowry et al. 2008). Studies using these measures focus primarily on organisms that can be held in the lab (e.g., *Drosophila,* plants and domestic species of birds). New methods to estimate reproductive isolation using phenotypic data from natural populations have recently been employed in the literature (Martin et al. 2010; Weir et al. 2012; Seddon et al. 2013). These methods use divergence in traits related to mate choice and species recognition – morphology, colour and song – as a proxy for reproductive isolation (Slabbekoorn and Smith 2002; Hill and McGraw 2006; Price 2008; Uy et al. 2009). There are two ways reproductive isolating barriers like seasonal migration could drive divergence in these traits: (1) through selection against maladaptive hybridization and competition (i.e., character displacement; Coyne and Orr 2004; Price 2008) and/or (2) as a byproduct of a reduction in gene flow caused by pre- and/or post-mating isolation which allows populations to follow different evolutionary trajectories.

In the present study, we set out to use divergence in phenotypic traits (morphology, colour and song) as a proxy for reproductive isolation to test the hypothesis that differences in migratory routes promote speciation in taxonomic groups that form migratory divides. We focused on birds because divides are common in this group (Irwin and Irwin 2005; Møller et al. 2011; Rohwer and Irwin 2011) and could be especially important for their speciation, as intrinsic postzygotic isolation is believed to play a limited role in avian speciation (Grant and Grant 1992; Price and Bouvier 2002; Rabosky and Matute 2013). We identified pairs of closely related bird

species or subspecies that breed in North America and compared phenotypic divergence between pairs that form divides and pairs that do not. We used mitochondrial sequence data to estimate time since divergence and predicted that the extent and rate of phenotypic divergence would be greater among pairs that form divides.

The geographic ranges of migratory species or subspecies in pairs that do not form divides are often restricted to either eastern or western North America. This is not the case in pairs that form divides; the range of one taxon is often restricted to the east and the other to the west (see Appendix D.1 for examples). This difference could result in greater secondary contact among pairs that do not form divides. Secondary contact can lead to either convergence or divergence in phenotypic traits (e.g., convergence can occur if species learn heterospecific acoustic signals, divergence can occur interspecific interactions lead to character displacement, Pfennig 1998; Haavie et al. 2004). We examined this potentially confounding effect on phenotypic divergence by including each pair's geographic distribution (allopatric, parapatric or sympatric) in our analyses.

In the end, our results were opposite to what we predicted; taxa that do not form divides exhibited higher apparent rates of phenotypic divergence than pairs that do. We explain this unexpected pattern as a possible result of the process of population fusion, with phenotypically similar groups being more likely to fuse together after secondary contact if they do not differ in migratory routes.

5.2 Methods

5.2.1 Study species

We identified 49 pairs of closely related migratory passerines and near-passerines that breed in North America using information from published molecular phylogenies (Lovette et al. 2010), literature reviews (Rohwer and Irwin 2011), focused studies of particular taxa (Boulet et al. 2006), Clements checklist (http://www.birds.cornell.edu/clementschecklist/), and Birds of North America (http://bna.birds.cornell.edu/bna/). We defined migrants as those that winter primarily south of the United States-Mexico border. 38/49 pairs were immediate sister taxa (i.e., each other's most closely related taxonomic group, including subspecies) and 11/49 pairs were from the same phylogenetic clade but not the most closely related group (e.g., if the most closely related group was not migratory, we chose the next most closely related migratory group). The latter set of pairs were never more than 2 branches apart and thus will still be relevant to our understanding of phenotypic divergence in recent evolutionary history (Seddon et al. 2013). We split these pairs into two categories based on whether they formed a migratory divide or not. Most pairs that form divides come into contact along the Rocky Mountains or the Great Plains. See Appendix D.1 for examples.

5.2.2 Measuring phenotypic traits

We measured divergence in three sets of phenotypic traits: morphology, plumage colour and song. To the best of our knowledge, all of the birds included in our study were breeding adult males (based on breeding plumage, location and date of capture [May–July, with a few from August for taxa with low sample sizes] for museum specimens and location and date of

recording [May–July] for recordings). When available, we included ten individuals per taxon (mean for colour and morphology = 9.89, range = 6–10; mean for song = 9.27, range = 4–10).

5.2.2.1 Morphology and colour

We measured morphology and colour from specimens at the Beaty Biodiversity Museum in Vancouver, British Columbia, and the Field Museum of Chicago in Chicago, Illinois. We measured six morphological variables: distance between the longest primary and secondary flight feathers, wing chord, tail, tarsus and bill length (all measured in mm, Pyle et al. 1997). We used an Ocean Optics USB2000 spectrometer (Dunedin, FL) to gather reflectance data from the feathers of seven body regions: crown, back, rump, tail, throat, belly, and wing coverts. We took five measurements from each region and plotted these data in tetrahedral colour space (Stoddard and Prum 2008) using the average avian UV visual model and the 'pavo' package (Maia et al. 2013) in R 3.0.3 (R Development Core Team 2014) to model the relative stimulation of each of the four cones in the avian visual system. In total, we collected measures of the maximum stimulation of each cone for all seven body regions, resulting in 28 plumage variables for each individual specimen.

5.2.2.2 Song

We focused our analysis on those vocalizations that play a major role in mate selection: songs of songbirds and territorial or pair-bonding calls in non-songbirds (Catchpole and Slater 2008). We obtained vocalizations from the Macaulay Library (http://macaulaylibrary.org), Xeno-canto (http://www.xeno-canto.org) and field recordings (KD, HK, DI). We used Syrinx version 2.6 (Burt 2005) to analyze these songs, measuring the following five variables from at least three

songs for each individual: duration (s), bandwidth (kHz), peak frequency (kHz), total number of syllables and number of syllable types. We defined a syllable as a unit within a vocalization which is made up of components that are always repeated together in the same sequence (Catchpole and Slater 2008).

5.2.3 Quantifying phenotypic divergence

We used Hedges' *g* to quantify divergence (Hedges 1981). This parameter is calculated by taking the difference between mean values for a given trait in two separate taxonomic groups and dividing it by the pooled standard deviation of the trait. Higher values of Hedges' *g* suggest that mean trait values are farther apart and/or that variability within each group is lower. We calculated Hedges' *g* for all variables and averaged these estimates for each set of traits separately (i.e., song, colour and morphology). We then averaged these summary values (i.e., averages for each set of traits) to obtain an overall estimate of divergence for each pair.

5.2.4 Measuring additional predictors of phenotypic divergence

5.2.4.1 Genetic distance

It can be expected that phenotypic divergence would tend to increase with the time that populations have evolved in separation. To control for this variation in divergence times, we estimated genetic distance between each pair using mitochondrial sequences from GenBank (https://www.ncbi.nlm.nih.gov/genbank). When available, we obtained sequence data from cytochrome b (available for 37 pairs). When data from this region were not available, we used other mitochondrial sequences (COI [9 pairs], control region [2 pairs], NADH [1 pair]) that have been shown to produce similar estimates of genetic distance (Weir and Wheatcroft 2011; Barker

et al. 2012). We aligned sequences using MEGA 6.06 and estimated genetic distances as p-distance, the proportion of nucleotide sites that differ between clades (Nei and Kumar 2000).

5.2.4.2 Geographic distribution, breeding latitude and migration distance

Secondary contact can result in phenotypic divergence or convergence. We accounted for this potentially confounding effect by including geographic distribution in our models. We downloaded range maps from Bird Life International and Nature Serve (Ridgely et al. 2011) and used ArcGIS (Esri, Redlands, California) to classify pairs as allopatric, parapatric and sympatric. We followed definitions from Price (2008) for these classifications, considering pairs allopatric if they were completely geographically separated, parapatric if they had abutting geographic ranges (some with hybrid zones) and sympatric if there was an absence of geographic separation over most of at least one taxon's range. Migratory divides are also more common at higher latitudes (Rohwer and Irwin 2011) and the strength of selection related to differences in migration may be greater for taxa that travel longer distances. We controlled for these effects by including estimates of breeding range latitude and migratory distance (averaged distance [km] between the center of each taxon's breeding and wintering ranges).

5.2.4.3 Body mass

We included body mass in our analyses to control for allometric effects on the traits measured. We obtained estimates of body mass from the CRC Handbook of Avian Body Masses (Dunning 1992). Where available, mean male mass was used, otherwise mass estimates come from both sexes or unknown sexes. For some pairs, mass estimates were not available at the subspecies level – for those the mean mass for the species was recorded for both members of the pair.

5.2.5 Statistical analysis

We used two complimentary approaches to evaluate the prediction that phenotypic divergence will be greater among taxa that form divides. First, we used an information theoretic approach with linear mixed effects models and model averaging (Burnham and Anderson 2002). We ran these analyses using the 'lmer', 'MuMin' and 'visreg' packages in R and included six fixed effects: migration category (divide or no divide), geographic distribution (coded as an ordinal variable with 0, 1 or 2 for allopatric, parapatric and sympatric, respectively), genetic distance, breeding range latitude, migration distance and body mass. We controlled for phylogenetic relationships in our dataset by including genus nested within family as a random effect in all models (Seddon et al. 2013). In total, we ran 64 models with average divergence as the response variable (all possible models from $n = 6$ predictor variables [2^n models, where n = number of predictors]). We selected models with ΔAICc (Akaike information criterion, corrected for sample size) less than 2 from the top model (Burnham and Anderson 2002) and ran these selected models again to recalculate relative AICc weights. Parameter estimates for each predictor were then averaged across the selected top models and weighted by the summed relative AICc weight of each model in which they were included (Burnham and Anderson 2002). This model averaging approach allows us to investigate how different combinations of the fixed effects influence phenotypic divergence, select the most parsimonious combination and understand how influential different migratory strategies are in relation to other predictors.

We also used Brownian motion (BM) and Ornstein-Uhlenbeck (OU) models to compare phenotypic divergence between different migratory categories. These models estimate rates of evolutionary divergence. BM models include a single parameter for evolutionary rate (β) and

assume that phenotypic divergence is unbounded. OU models extend BM models by including an additional parameter (α) that constrains phenotypic divergence to an intermediate value, constraining phenotypic divergence. The constraint parameter (α) varies from 0 to 1, with higher values making divergence more difficult. We ran these analyses using the 'EvoRAG' package in R and compared models with the same or separate evolutionary rates for each migratory category. We considered the model with the lowest AICc value the best fit to the data (Burnham and Anderson 2002). These models allow us to examine rates of phenotypic divergence in more detail than our model averaging approach, including a comparison between models with and without evolutionary constraint.

5.3 Results

Phenotypic divergence between pairs ranged from 0.44 to 3.93 (Hedges' *g* values; mean = 1.37). To examine the contribution of each trait set (i.e., song, colour, morphology) to total divergence, we identified the trait set with the highest Hedges' *g* value in each pair. The number of pairs maximally diverged in the song, colour and morphological datasets was not different than expected by chance (18 morphology, 18 plumage, 16 song; expected values = 16.33; $\chi^2 = 0.124$, $df = 2$, $P = 0.93$).

Using the model averaging approach to investigate how different combinations of the fixed effects influenced overall phenotypic divergence, we found that migration category (divide or no divide), time since divergence and body mass were significant predictors of overall phenotypic divergence (i.e., confidence intervals on estimates did not overlap zero, Table 5.1a). The modelled relationship between migration category and phenotypic divergence is shown in Figure 5.1 (relationships for all variables included as predictors in the global model are shown in

Appendix D.2). Contrary to our predictions, phenotypic divergence was greater among pairs that do not form divides than among those that do. Time since divergence and body mass had positive effects on phenotypic divergence (Table 5.1a).

The results described above considered all traits together (i.e., average divergence in song, colour and morphological traits). We also used model averaging to examine the relationship between our predictor variables and divergence in each set of traits separately. Migration category and time since divergence emerged as significant predictors in the average best model for divergence in song and colour (Table 5.1bc). Time since divergence and breeding latitude were the only significant predictors of morphological divergence (Table 5.1d). Similar to results using divergence in all trait sets, pairs that formed migratory divides had lower values of song and colour divergence and time since divergence had a positive effect on divergence in both traits. Geographic distribution (allopatric [0], parapatric [1], or sympatric [2]) did not emerge as a significant predictor in the average best model of any trait set (separate or combined, Table 5.1).

Figure 2 shows the relationship between phenotypic divergence and time since divergence for all traits sets. In each case, the evolution of phenotypic divergence appears to be bounded, with divergence occurring rapidly at first and slowing down with increased time since divergence. ΔAICc values from a comparison of BM and OU models support this observation, with OU models providing the best fit to the data in all comparisons (Table 5.2; Figure 5.2). The best fit model in all cases (when all traits were combined and when each trait set was examined separately) included separate rates for each migration category. Similar to results from model averaging and contrary to our original prediction, rates of evolution (β) were lower for pairs that form divides (Table 5.2; Figure 5.2).

To examine the effect of geographic distribution on our results, we reran the best fit OU models (i.e., those allowing separate rates for migration category) with and without geographic distribution. These models did not provide a better fit to the data (all log-likelihoods were within 1.92 units of models without geographic distribution, Weir et al. 2010). In addition, while the average rate of phenotypic divergence generally increased with geographic distribution in these models, 95% confidence intervals around these rates overlapped 0 in all cases, suggesting this increase was non-significant. Appendix D.3 includes results from these models and Appendix D.4 shows the relationships from Figure 5.2 along with the geographic distribution of each pair.

5.4 Discussion

In this study we used a quantitative and comparative approach to examine a prediction based on a prominent hypothesis in the speciation literature, that differences in the migratory routes of populations that form migratory divides promote reproductive isolation (Helbig 1991; Bearhop et al. 2005; Irwin and Irwin 2005; Veen et al. 2007; Bensch et al. 2009; Rolshausen et al. 2009; Møller et al. 2011; Rohwer and Irwin 2011; Delmore and Irwin 2014). Our broad statistical analysis of divergence between migratory sister pairs showed that the extent and rate of phenotypic divergence was greater for pairs that do not form migratory divides. These results are opposite to what we predicted; we reasoned that, if phenotypic divergence is a proxy for reproductive isolation and differences in migration promote speciation, pairs that form divides should exhibit higher rates of divergence. We offer an interpretation of these findings below that invokes the geographic history of North American birds and the process of differential fusion.

Sister pairs of bird taxa likely have complex histories that involved periods of both allopatry and secondary contact (Hewitt 2000; Weir and Schluter 2004; Lovette 2005). Most

pairs of taxa have experienced many cycles of glacial advances and retreats since their original divergence from a common ancestor. During periods of glacial advance, species were divided into allopatric populations, but during interglacial periods these populations would likely have expanded their ranges and come into secondary contact. Gene flow following secondary contact would likely have resulted in the fusion of pairs that had not diverged sufficiently in traits involved in reproductive isolation (Arnold 1997). There are a wide variety of traits that could have prevented fusion upon secondary contact, including migratory differences and the other phenotypic traits examined here. If migration does play a strong role in reproductive isolation, then the process of fusion would be less likely to occur in pairs with migratory divides. By contrast, pairs without migratory divides that do not differ in other phenotypic traits would be likely to fuse, such that they would not be included in a study of sister pairs, whereas those that do differ strongly in song, color, or morphometrics would be more likely to remain distinct forms. Altogether, this process of differential fusion (Templeton 1981; Noor 1999) would lead to the observed pattern, of greater observed phenotypic difference in pairs without migratory divides.

The interpretation above suggests that differences in the role of migration as a reproductive isolating barrier between pairs that do and do not form divides explains increased apparent rates of divergence in the latter group. As noted in the Introduction, sister pairs that do not form divides are often found either in western or eastern North America. By contrast, pairs that form divides often include one taxon in the west and the other in the east (Appendix D.1). Accordingly, pairs that do not form divides are more proximate and may, as a result, have come into secondary contact more often than those that do form divides. Secondary contact can lead to increased phenotypic divergence through character displacement and/or differential fusion.

Character displacement is an increase in phenotypic differences driven by selection to reduce resource competition or reproductive interactions (Pfennig and Pfennig 2010). Differential fusion describes the apparent increase in mean rates of phenotypic divergence that results from the selective elimination of pairs that were not diverged sufficiently in allopatry to remain distinct (described above). Accordingly, increased secondary contact among pairs that do not form divides might account at least partially for increased rates of divergence in this group. Nevertheless, the bulk of evidence presented in our study does not support this suggestion: geographic distribution was not included in any of the final averaged linear mixed effects models and models of trait evolution that included geographic distribution did not provide a better fit to the data than those without this variable. In addition, while there was a general increase in divergence with geographic distribution in models of trait evolution, these increases were not significant.

Instead, several additional details support the first interpretation of our results, that secondary contact and differences in the main reproductive isolating barriers between migratory categories account for increased phenotypic divergence in those that do not form divides. To begin with, following this interpretation we would expect to find an excess of pairs with low values of trait divergence in the divide category. This is exactly what we see in Figure 5.2, where there are more young pairs with low differentiation in the divide group (open circles, bottom left hand corner). Sister pairs that form divides today also provide support for this expectation, as many of the migratory divides described to date occur between taxa that exhibit low levels of phenotypic divergence, including subspecies of willow warblers (Bensch et al. 2009), populations of European blackcaps (Pérez-Tris et al. 2004) and subspecies groups of the Swainson's thrush. For example, two groups of Swainson's thrushes form a migratory divide in

western North America, where they differ in song, morphology and colour (Ruegg et al. 2006b; Ruegg 2008). Differences in migratory routes distinguish these groups more clearly than the other traits (Cohen's *D* of 0.01 for migration, 3.61 for song, 2.15 for morphology and 0.28 for colour; Delmore and Irwin 2014). Finally, a comparison between migratory and non-migratory taxa supports the suggestion that migration could be maintaining reproductive isolation between taxa that form divides. Specifically, as suggested above, most migratory divides comprise one taxon in the east and another in the west. This is not the case for non-migratory groups like the red-breasted nuthatch, brown creeper, ruby- and golden-crowned kinglets, black-capped chickadee, American robin and crow; most of these groups have ranges that span the entire continent. During the glacial advances these species would likely have experienced similar periods of allopatry to taxa that form divides, occupying both eastern and western refugia. Accordingly, they might be expected to exhibit a similar east/west division. Perhaps the fact that they did not differ in migratory behaviour prevented them from maintaining differences they evolved in allopatry.

Future analyses could take a variety of approaches to expand on our results. For instance, whereas we used divergence in phenotypic traits as our measure of reproductive isolation, studies using a more direct measure of this variable would be valuable. Data on hybrid zone width is often used to examine the strength of reproductive isolation, with narrower hybrid zones suggesting stronger reproductive isolation (Barton and Hewitt 1985). Using a measure of reproductive isolation that is not also involved in generating reproductive isolation would allow us to more directly evaluate the prediction that migratory divides promote speciation in birds. At this time, data for the width of 13 hybrid zones between pairs in our study are available and suggest that hybrid zones between taxa that form divides are in fact narrower than hybrid zones

between taxa that do not (no divide, mean = 289 km, range = 32 – 600; divide, mean = 163 km, range = 60 – 400), supporting a role for migration in reproductive isolation. With the current surge of next-generation sequencing, it is likely that additional data on hybrid zones will become available, allowing a more quantitative approach to this analysis. One additional extension that would be informative is to collect similar data in other migratory systems. The main ecological barriers encountered by birds breeding in North America are mountain chains and deserts in western North America. These barriers may only moderately reduce the fitness of hybrids. By contrast, birds breeding in Siberia and encountering the Tibetan Plateau and Gobi desert on their way to Asia may experience greater reductions in fitness on migration than birds breeding in North America (Irwin and Irwin 2005).

The results we presented here are counter to our original prediction, that taxa with migratory divides would exhibit higher levels of trait divergence. Nevertheless, migratory category was a significant predictor of divergence in most analyses, with taxa that do not form divides being more diverged than those that do. This finding suggests that migration or some correlated factor has a strong effect on observed phenotypic divergence between closely related groups of North American birds. We have outlined one interpretation of these results, that sister pairs included in our study have come into secondary contact throughout their geographic histories and those that were not sufficiently differentiated in one or more traits involved in reproductive isolation have fused into a single unit. These traits were related to migration for pairs that form divides and other phenotypic traits like those we measured in our study for pairs that do not form divides. Additional work (e.g., analyzing data from hybrid zones) will be needed to thoroughly test this suggestion.

Table 5.1 Results from averaged linear mixed models examining the relationship between phenotypic divergence and a series of predictor variables, including migration category (divide or no divide), time since divergence (p-distance, the proportion of nucleotide sites that differ between clades), geographic distribution (ordinal variable from allopatric, parapatric to sympatric), migration distance, breeding latitude and body mass. Models were run with family and genus as nested random variables and total divergence as a dependent variable (for all traits combined [a] and each set of traits separately [b-d]). Only predictor variables included in the final averaged model are shown; those that are significant predictors of phenotypic divergence (where confidence intervals [CI] do not overlap zero) are shown in bold. Negative parameter estimates for migration category indicate that groups with migratory divides tend to have lower phenotypic divergence than groups without divides. n = number of models in global model (i.e., within ΔAICc of 2), R^2 = goodness-of-fit of the global model.

Predictor Variables	Parameter Estimates	Standard Errors	Lower CI	Upper CI
a) Overall phenotypic divergence (R^2 = 0.57, n = 4)				
(Intercept)	1.73	1.27	-0.79	4.25
Migration category	**-0.44**	**0.11**	**-0.65**	**-0.22**
Time since divergence	**0.22**	**0.047**	**0.13**	**0.32**
Breeding latitude	-0.51	0.35	-1.22	0.20
Body mass	**0.23**	**0.098**	**0.03**	**0.43**
b) Song divergence (R^2 = 0.36, n = 5)				
(Intercept)	-0.19	1.62	-3.41	3.039
Migration category	**-0.57**	**0.16**	**-0.89**	**-0.25**
Time since divergence	**0.18**	**0.070**	**0.039**	**0.32**
Breeding latitude	0.43	0.53	-0.63	1.50
Migration distance	0.30	0.17	-0.036	0.63
c) Colour divergence (R^2 = 0.40, n = 3)				
(Intercept)	1.65	1.52	-1.38	4.70
Migration category	**-0.46**	**0.17**	**-0.80**	**-0.11**
Time since divergence	**0.27**	**0.076**	**0.11**	**0.42**
Breeding latitude	-0.50	0.59	-1.68	0.68
Body mass	0.32	0.17	-0.023	0.66
d) Morphological divergence (R^2 = 0.30, n = 2)				
(Intercept)	5.75	2.12	1.48	10.03
Time since divergence	**0.19**	**0.072**	**0.045**	**0.34**
Breeding latitude	**-1.17**	**0.56**	**-2.47**	**0.25**

Table 5.2 Results from Brownian Motion and Ornstein-Ulhenbeck (OU) models of evolutionary change. Models including a single evolutionary rate are compared with models fitting separate rates for each migration category (no divide, divide). Models were run with all traits combined (a) and each set of traits separately (b to d). The model with the lowest ΔAICc value is shown in bold. OU models include and additional parameter (α) that constrains the rate (β) of divergence to an intermediate value.

	Brownian motion		Ornstein-Uhlenback		
	β	ΔAICc	β	α	ΔAICc
a) Overall phenotypic divergence					
one rate	0.0061	5.28	0.018	0.62	3.83
no divide/divide	0.0097/0.0038	7.00	0.040/0.006	1.18/0.18	**0**
b) Song divergence					
one rate	0.0053	8.42	0.013	0.48	9.48
no divide/divide	0.0092/0.0028	8.63	0.021/0.0085	0.42/0.69	**0**
c) Colour divergence					
one rate	0.0087	5.68	0.020	0.43	4.69
no divide/divide	0.014/0.0053	5.23	0.037/0.0092	0.57/0.24	**0**
d) Morphological divergence					
one rate	0.0078	2.87	0.021	0.55	9.69
no divide/divide	0.011/0.0056	13.09	0.014/0.0059	6.43/0.018	**0**

Figure 5.1 Modelled relationship between overall phenotypic divergence and migration category (no divide, divide). Model averaging with linear mixed effects models were run with a random variable of genus nested within family to control for phylogenetic relationships. Significant predictors in the final model included migration category, time since divergence and body mass. The effect of migration category is plotted with these additional variables held at their medians.

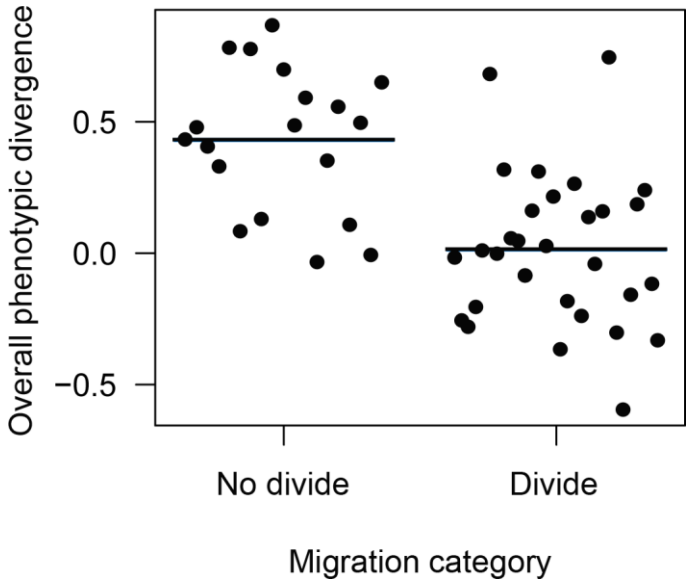
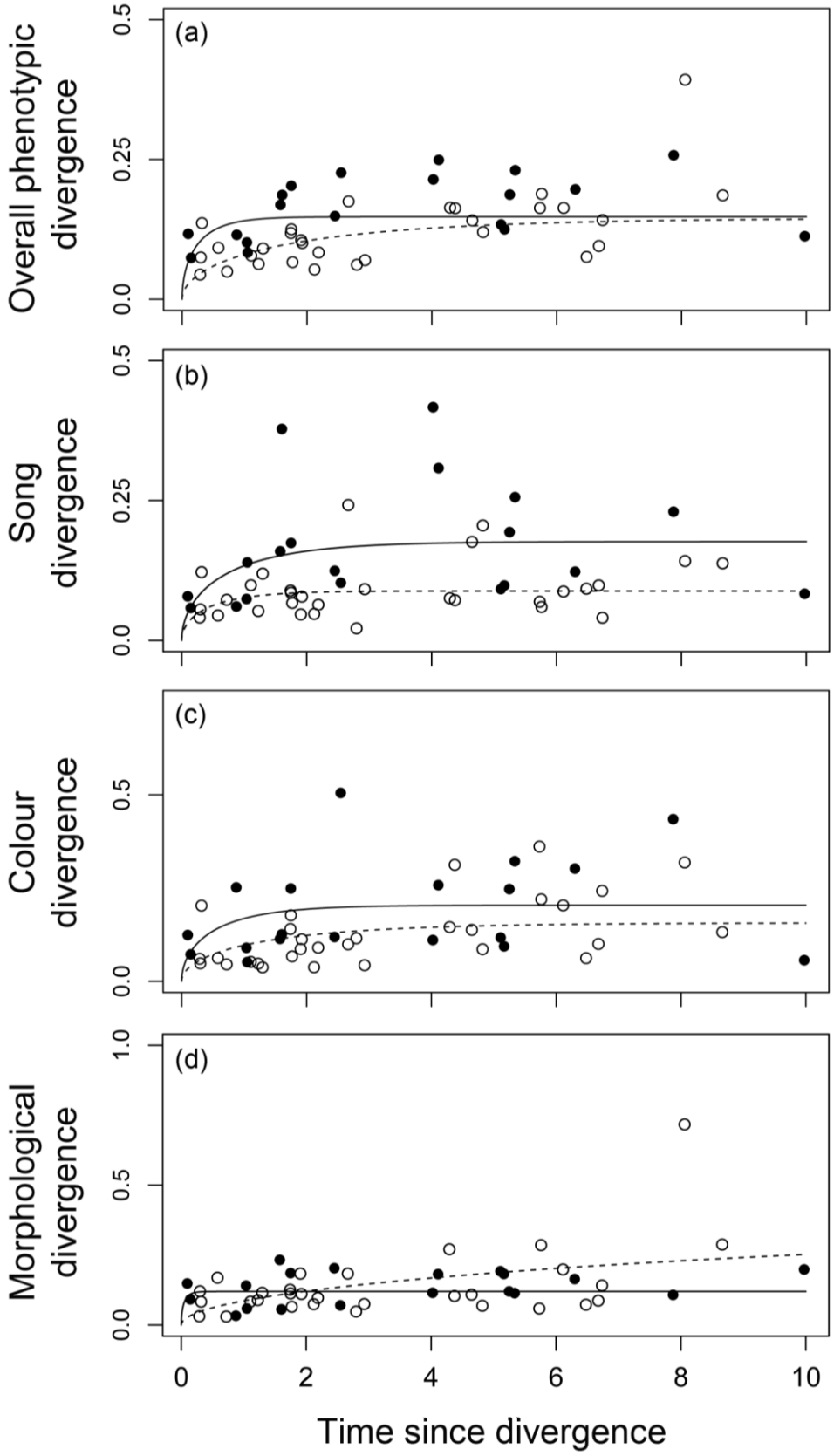


Figure 5.2 Relationship between phenotypic divergence and time since divergence. Results are shown for all traits combined (a) and each set of traits separately (b to d). Time since divergence was measured using genetic distance; phenotypic divergence was measured using average Hedges' *g*. Curves show results from OU models of evolutionary change, fitting separate models for taxa that do (dotted line and open points) and do not (black line and filled points) form migratory divides.


(a)
(b)
(c)
(d)
Overall phenotypic divergence
Song divergence
Colour divergence
Morphological divergence
0.0
0.25
0.5
1.0
0
2
4
6
8
10
Time since divergence

Chapter 6: General discussion

6.1 Migration's role as a reproductive isolating barrier between Swainson's thrushes

The first three chapters of my dissertation focused on the Swainson's thrush. I began this work by using tracking data from birds caught at the edges of the hybrid zone between two subspecies groups of this species to confirm previous findings from band recovery data and genetic markers suggesting that these groups form a migratory divide (Ruegg and Smith 2002; Chapter 2). Once the divide was confirmed, I used additional tracking data from birds caught at the center of the hybrid zone to evaluate the main prediction from the migratory divide hypothesis, that hybrids take intermediate routes. My results supported this prediction, with many hybrids exhibiting intermediate behaviours on migration; some of these birds combined parental routes to fly midway between parental forms while others use an approximately inland route on fall migration and coastal route on spring migration (Chapter 3).

The migratory divide hypothesis also predicts that the intermediate routes taken by hybrids will be inferior. I obtained estimates of selection against hybrids using cline theory, and found that selection against this group was high, especially compared to estimates from other hybrid zones (Chapter 3). Several lines of evidence suggest that differences in migratory orientation contribute to this selection. For example, birds who took routes midway between parental forms migrated over several barriers in the southwestern United States that are likely difficult to navigate and provide little habitat to forage (Chapter 3). I also documented substantial genomic differentiation between inland and coastal thrushes. Candidate genes for migration were enriched in regions showing the highest levels of differentiation and other signatures of selection

(Chapter 4, see below), providing further evidence for the role migration could play in reducing gene flow between Swainson's thrushes.

Together, this work represents the most comprehensive and direct evaluation of the migratory divide hypothesis conducted to date. Future research to more directly evaluate the second prediction from the migratory divide hypothesis, that hybrid routes are inferior, could compliment this work well. One immediate option could be to combine niche modelling with band recovery data from the migratory period (April-May and September-October) to identify the most suitable niche axes for Swainson's thrushes on migration. A comparison of habitat quality along parental vs. hybrid routes would help determine whether the hybrid routes are less suitable for migration. These data could also be combined with a connectivity analysis (e.g., least-cost path analyses and/or circuit analyses; Adriaensen et al. 2003; McRae et al. 2008) to identify the most suitable routes through North America for Swainson's thrushes on migration.

Future advances in tracking technology could also advance to this work. Light-level geolocators are the only technology currently available to track songbirds on migration. A new generation of geolocators fitted with radio transmitters are currently in development and will make it easier to recapture birds the following year (Lotek, pers. comm.). This addition could allow us to track juvenile birds on migration; this was not an option before, as juveniles rarely return to their natal grounds to breed, making it nearly impossible to find them the next year. The genetic component of migration is strongest in juveniles (Berthold 1996) and thus it would be interesting to compare the behaviour of these birds to adults. A much more transformative development would be the miniaturization of devices that provide daily estimates of location via satellites, as these devices would allow us to track unsuccessful migrants as well. We could determine where these birds stop moving and in some systems even retrieve them. This

development is on the horizon, with ICARUS (International Cooperation for Animal Research Using Space) and the Max Planck Institute for Ornithology deploying the first satellite tags on songbirds this summer.

6.2 Genetic basis of reproductive isolation in Swainson's thrushes

Results from Chapter 4 lent support to migration's role as a reproductive isolating barrier between Swainson's thrushes but are more generally relevant to the second objective of my PhD, which was to begin studying the genetic basis of reproductive isolation between these subspecies groups. Using whole genome sequence (WGS) data from populations at the edges of the hybrid zone between Swainson's thrushes, I documented genome-wide heterogeneity in genetic differentiation.

Some of the most prominent voices in the literature have argued that differential gene flow and divergent selection may be generating genomic heterogeneity in many systems (Nosil et al. 2009; Feder et al. 2012; Via 2012). My results suggest that the Swainson's thrush is not one of these systems. Instead, selective sweeps within populations and variation in recombination rates across the genome likely generate these patterns. To begin with, reductions in both within population variation (π) and absolute genetic differentiation (d_{XY}) occurred in areas of high relative differentiation between thrushes (F_{ST}). These reductions were often located near centromeres and the center of chromosomes (e.g., chromosome 4, Figure 4.2). Recombination is lower in both these areas (even in non-metacentric chromosomes, Backström et al. 2010; Roesti et al. 2013). It should also be noted, that Ruegg et al. (2014) used RAD tags to estimate genomic differentiation between allopatric populations of inland and coastal thrushes far from the contact zone. They found a similarly heterogeneous landscape of differentiation. This would not be

expected if gene flow were generating these patterns, as differentiation should be more homogeneous in allopatric comparisons (Nosil et al. 2009).

The finding that selective sweeps and recombination generate genomic heterogeneity in the thrush is quite timely, as there is currently a switch in the literature to acknowledge the role these forces (vs. gene flow) could be playing (Cruickshank and Hahn 2014). Nevertheless, my work has only begun to scratch the surface of this question. Future research will be aided by the infrastructure developed here, including the identification of genomic differences between parental populations using WGS data and assembly of a draft reference genome for this species. Models of demographic history using these data could be particularly informative; phylogenetic studies suggest that Swainson's thrushes diverged during the Pleistocene (Ruegg and Smith 2002). They likely underwent periods of allopatry and secondary contact throughout this period with several opportunities to exchange genes (Hewitt 2000; Lovette 2005). Genomic data obtained during my dissertation could be used to determine if gene flow occurred during these periods. Several new approaches have been developed for these analyses (summarized in Sousa & Hey 2013) and it should become possible to identify specific regions of the genome that have experienced more or less gene flow.

Swainson's thrushes form a hybrid zone in the Coast Mountains. Data from hybrids could be used to implicate the candidate genes I found in areas of elevated differentiation more directly in speciation and migration. For instance, we could combine genomic data from hybrids with information on their orientation (e.g., from geolocators and/or orientation cages) to conduct admixture mapping, identifying genomic regions associated with migratory orientation (Rieseberg and Buerkle 2002; Buerkle and Lexer 2008). Once these regions are identified, we could determine whether they are located in areas of increased differentiation between parental

forms. We could also use geographic (Barton and Hewitt 1985) and/or genomic (Gompert and Buerkle 2011) clines to determine whether these regions experience less introgression than the rest of the genome. This area of research would be particularly interesting, as researchers have only recently started making progress towards identifying the genes associated with migration (Mueller et al. 2011) and there has been no progress to date in identifying loci associated with orientation specifically (vs. other migratory traits like the propensity to migrate).

6.3 Expanding out from the Swainson's thrush

The hybrid zone between Swainson's thrushes occurs in a suture zone (i.e., an area where many boundaries between neighbouring taxa [species, subspecies, or phylogroups] co-occur; Swenson and Howard 2004). Taxa at suture zones likely experience similar geographic histories. Accordingly, it may be possible to extrapolate from the Swainson's thrush to other systems. I set out to evaluate this suggestion in Chapter 5, by examining levels of phenotypic divergence between pairs of closely related North American bird taxa, with the prediction that divergence would tend to be greater in pairs that form migratory divides than in pairs that do not form divides. The main assumption of this work was that phenotypic divergence would be greater among taxa that were more reproductively isolated, as these traits are involved in species recognition (Martin et al. 2010; Tobias et al. 2010; Weir and Wheatcroft 2011).

Contrary to my original prediction, phenotypic divergence was not greater among taxa that form divides; in fact, divergence was greater in those that do not form divides. This might at first seem to suggest that migration does not promote reproductive isolation; perhaps the Swainson's thrush is a special case. The Swainson's thrush does migrate further than many other North American taxa (e.g., many wood warblers [*Setophaga*] do not migrate south of Mexico),

and thus encounters more geographic barriers that could reduce the fitness of hybrids. The Swainson's thrush is also an understory bird that spends its time on the breeding grounds in dense, wet forests (Mack et al. 2000). Accordingly, this species may be more affected by the open dry habitat of the southwestern United States. Even if this were the case, the observed increase in divergence documented between taxa that do not form divides needs to be explained.

I believe that a combination of North America's geographic history and migration may explain this pattern. As mentioned above, populations in North America have been in flux throughout the Pleistocene, experiencing periods of allopatry and secondary contact. Upon secondary contact, closely related groups would have fused into a single unit unless they had diverged enough in allopatry to prevent homogenization following gene flow. For pairs that form divides, divergence in migration may have been the key to preventing their fusion. Pairs that do not form divides may instead be those that diverged in the phenotypic traits we measured, which have been shown to play a role in speciation (Coyne and Orr 2004; Price 2008). This process would have permitted the existence of taxa in divides that exhibit lower levels of phenotypic divergence. The Swainson's thrush typifies this idea, with limited divergence in phenotypic traits is observed between inland and coastal birds (Chapter 3).

6.4 Broader implications

6.4.1 Speciation

I outlined several avenues for future research into migration's role in reducing gene flow between Swainson's thrushes above, including habitat connectivity analyses and the use of newer tracking technologies. I strongly believe that this is a worthwhile pursuit, not only in the Swainson's thrush but other systems as well; following my interpretation of results in Chapter 5,

migration is among the major sources of reproductive isolation between North American birds and additional divides have been documented worldwide (Irwin and Irwin 2005; Møller et al. 2011). Intrinsic postzygotic barriers likely play a limited role in birds (Price and Bouvier 2002; Rabosky and Matute 2013) and the number of hybrids documented in nature (Grant and Grant 1992; McCarthy 2006) suggests that premating isolation is rarely complete in this group. Migration could be among the extrinsic postzygotic barriers that fill this gap. It should be noted, that I have focused on migratory orientation here but differences in additional migratory traits likely also play a role and warrant further research, including differences in arrival time (Bearhop et al. 2005; Friesen et al. 2007) and molting schedules (Rohwer et al. 2005).

Results from my dissertation are not only relevant to birds, as members of almost every animal group migrate and migratory divides have been identified in mammals (e.g., O'Corry-Crowe et al. 1997), insects (Altizer and Davis 2010) and fishes (Gagnaire et al. 2009). My results are also in line with the current sentiment that ecology plays an important role in speciation (Schluter 2000; Rundle and Nosil 2005; Nosil 2012) and are on point for researchers interested in the role behavioural traits could play in this process (Hoekstra 2010). In particular, there is a good deal of interest in studying the genetic basis of behavioural traits involved in speciation, as much of the research in this area has focused on morphological traits (e.g., beak size in finches, Grant & Grant 1993; body armour in sticklebacks, Colosimo et al. 2005). This focus has made it difficult to make generalizations about the genetic basis of speciation. Challenges involved in quantifying behaviours are likely responsible for this limitation. Combined with new tracking technologies, migratory divides could represent an ideal system for this work. The project I described above, using admixture mapping and cline analyses would be a great start for this work.

6.4.2 Conservation

It is difficult to work on such a recognizable and ecologically important group of organisms without considering how your work relates to their conservation. Migratory songbirds are currently experiencing drastic population declines in response to climate change and habitat loss. These birds spend more than half the year away from their breeding grounds and yet this is the time for which we have the least data (Faaborg et al. 2010b; Hobson et al. 2014; Hostetler et al. 2015). The advent of geolocators has begun to change this and results from my work provide an excellent example of their potential value. For instance, I documented strong connectivity within subspecies groups of the Swainson's thrush and identified important stopover sites for each group. Other songbirds likely use similar routes and thus this information can be used to establish chains of stopover sites to focus management efforts for species in decline. Geolocators are being deployed on many different species now, and I anticipate that great advances in our understanding of the non-breeding ecology of migrants will allow us to establish more detailed conservation strategies for this group.

Gene flow over the geographic history of these populations was a common theme in my work. Despite this fact, inland and coastal Swainson's thrushes appear to have maintained their differences, as background levels of genomic differentiation were quite high between these groups. In fact, occasional gene flow may have actually worked in their favour, by allowing advantageous mutations that evolved in allopatry to pass between them. It is possible that this process will help Swainson's thrushes and other groups with similar geographic histories cope in our continually modified contemporary world. For example, climate change is expected to change migration routes (Faaborg et al. 2010b). Gene flow could provide these species with the

variation they need to adapt to these changes, especially those that exhibit low levels of divergence and continue to exchange genes now. This is one, among many instances, where hybridization may serve as a creative force in evolution (Dowling and Secor 1997; Mallet 2007; Schwenk et al. 2008) and further work on this topic and those outlined above will undoubtedly uncover many additional, exciting avenues of research for migratory divides.

References

Adriaensen, F., J. P. Chardon, G. De Blust, E. Swinnen, S. Villalba, H. Gulinck, and E. Matthysen. 2003. The application of "least-cost"modelling as a functional landscape model. Landscape and Urban Planning 64:233–247.

Albert, A. Y. K., and S. P. Otto. 2005. Sexual selection can resolve sex-linked sexual antagonism. Science 310:119–121.

Alerstam, T. 2006. Conflicting evidence about long-distance animal navigation. Science 313:791–794.

Alerstam, T. 2001. Detours in bird migration. Journal of Theoretical Biology 209:319–331.

Altizer, S., and A. K. Davis. 2010. Populations of monarch butterflies with different migratory behaviors show divergence in wing morphology. Evolution 64:1018–1028.

Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. Journal of Molecular Biology 215:403–410.

Andrew, R. L., K. L. Ostevik, D. P. Ebert, and L. H. Rieseberg. 2012. Adaptation with gene flow across the landscape in a dune sunflower. Molecular Ecology 21:2078–2091.

Arnold, M. 1997. Natural hybridization and evolution. Oxford University Press, New York.

Bächler, E., S. Hahn, M. Schaub, R. Arlettaz, L. Jenni, J. W. Fox, V. Afanasyev, and F. Liechti. 2010. Year-round tracking of small trans-Saharan migrants using light-level geolocators. PLoS ONE 5:e9566.

Backström, N., W. Forstmeier, H. Schielzeth, H. Mellenius, K. Nam, E. Bolund, M. T. Webster, T. Öst, M. Schneider, B. Kempenaers, and H. Ellegren. 2010. The recombination landscape of the zebra finch *Taeniopygia guttata* genome. Genome Research 20:485–495.

Backström, N., N. Karaiskou, E. H. Leder, L. Gustafsson, C. R. Primmer, A. Qvarnström, and H. Ellegren. 2008. A gene-based genetic linkage map of the collared flycatcher (*Ficedula albicollis*) reveals extensive synteny and gene-order conservation during 100 million years of avian evolution. Genetics 179:1479–1495.

Backström, N., A. Qvarnström, L. Gustafsson, and H. Ellegren. 2006. Levels of linkage disequilibrium in a wild bird population. Biology Letters 2:435–438.

Barker, F. K., M. K. Benesh, A. J. Vandergon, and S. M. Lanyon. 2012. Contrasting evolutionary dynamics and information content of the avian mitochondrial control region and ND2 gene. PloS ONE 7:e46403.

Bartell, P. A., and E. Gwinner. 2005. A separate circadian oscillator controls nocturnal migratory restlessness in the songbird *Sylvia borin*. Journal of Biological Rhythms 20:538–549.

Barton, N. H., and K. S. Gale. 1993. Genetic analysis of hybrid zones. Pp. 13–45 *in* R. G. Harrison, ed. Hybrid zones and the evolutionary process. Oxford University Press, New York.

Barton, N. H., and G. M. Hewitt. 1989. Adaptation, speciation and hybrid zones. Nature 341:497-503.

Barton, N. H., and G. M. Hewitt. 1985. Analysis of hybrid zones. Annual Review of Ecology and Systematics 16:113–148.

Bearhop, S., W. Fiedler, R. W. Furness, S. C. Votier, S. Waldron, J. Newton, G. J. Bowen, P. Berthold, and K. Farnsworth. 2005. Assortative mating as a mechanism for rapid evolution of a migratory divide. Science 310:502–504.

Bensch, S., T. Andersson, and S. Åkesson. 1999. Morphological and molecular variation across a migratory divide in willow warblers, *Phylloscopus trochilus*. Evolution 53:1925–1935.

Bensch, S., M. Grahn, N. Müller, L. Gay, and S. Åkesson. 2009. Genetic, morphological, and feather isotope variation of migratory willow warblers show gradual divergence in a ring. Molecular Ecology 18:3087–3096.

Berthold, P. 2001. Bird migration: a general survey. Oxford University Press, New York.

Berthold, P. 1996. Control of bird migration. Chapman and Hall, London.

Berthold, P. 1991. Spatiotemporal programmes and genetics of orientation. Pp. 86–105 *in* P. Berthold, ed. Orientation in birds. Birkhäuser, Basel.

Blanco, E., G. Parra, and R. Guigó. 2007. Using geneid to identify genes. Current Protocols in Bioinformatics Chapter 4: Unit 4–3.

Boulet, M., H. L. Gibbs, and K. A. Hobson. 2006. Integrated analysis of genetic, stable isotope, and banding data reveal migratory connectivity and flyways in the northern yellow warbler (*Dendroica petechia*; Aestiva group). Ornithological Monographs 61:29–78.

Brelsford, A., and D. E. Irwin. 2009. Incipient speciation despite little assortative mating: the yellow-rumped warbler hybrid zone. Evolution 63:3050–3060.

Buerkle, C. A., and C. Lexer. 2008. Admixture as the basis for genetic mapping. Trends in Ecology & Evolution 23:686–694.

Burnham, K. P., and D. R. Anderson. 2002. Model selection and multimodel inference: a practical information-theoretic approach. Springer, New York.

Burt, J. 2005. Syrinx, version 2.6. University of Washington, Seattle, Washington (available at www. syrinxpc. com).

Cantarel, B. L., I. Korf, S. M. C. Robb, G. Parra, E. Ross, B. Moore, C. Holt, A. S. Alvarado, and M. Yandell. 2008. MAKER: An easy-to-use annotation pipeline designed for emerging model organism genomes. Genome Research 18:188–196.

Carling, M. D., and R. T. Brumfield. 2009. Speciation in *Passerina* buntings: introgression patterns of sex-linked loci identify a candidate gene region for reproductive isolation. Molecular Ecology 18:834–847.

Catchpole, C. K., and P. J. B. Slater. 2008. Bird song: biological themes and variations. Cambridge University Press, New York.

Chamberlain, C., S. Bensch, X. Feng, S. Akesson, and T. Andersson. 2000. Stable isotopes examined across a migratory divide in Scandinavian willow warblers (*Phylloscopus trochilus trochilus* and *Phylloscopus trochilus acredula*) reflect their African winter quarters. Proceedings of the Royal Society of London. Series B: Biological Sciences 267:43–48.

Cochran, W. W., M. Wikelski, R. Greenberg, and P. P. Marra. 2005. Individual migratory tactics of New World *Catharus* thrushes. Pp. 13-45 *in* R. Greenberg and P. P. Marra, eds. Birds of two worlds: the ecology and evolution of migration. Johns Hopkins University Press, Baltimore.

Colosimo, P. F., K. E. Hosemann, S. Balabhadra, G. Villarreal, M. Dickson, J. Grimwood, J. Schmutz, R. M. Myers, D. Schluter, and D. M. Kingsley. 2005. Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. Science 307:1928–1933.

Coyne, J. A., and H. A. Orr. 1997. "Patterns of speciation in *Drosophila*" revisited. Evolution 295–303.

Coyne, J. A., and H. A. Orr. 2004. Speciation. Sinauer Associates, Massachusetts.

Cruickshank, T. E., and M. W. Hahn. 2014. Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. Molecular Ecology 23:3133–3157.

Cuevas, J. M., P. Domingo-Calap, and R. Sanjuán. 2012. The fitness effects of synonymous mutations in DNA and RNA viruses. Molecular Biology and Evolution 29:17–20.

Dallimer, M., and P. J. Jones. 2002. Migration orientation behaviour of the red-billed quelea *Quelea quelea*. Journal of Avian Biology 33:89–94.

Dallimer, M., P. J. Jones, J. M. Pemberton, and R. A. Cheke. 2003. Lack of genetic and plumage differentiation in the red-billed quelea *Quelea quelea* across a migratory divide in southern Africa. Molecular Ecology 12:345–353.

Davey, J. W., P. A. Hohenlohe, P. D. Etter, J. Q. Boone, J. M. Catchen, and M. L. Blaxter. 2011. Genome-wide genetic marker discovery and genotyping using next-generation sequencing. Nature Reviews Genetics 12:499–510.

Davis, L. A., E. H. Roalson, K. L. Cornell, K. D. Mcclanahan, and M. S. Webster. 2006. Genetic divergence and migration patterns in a North American passerine bird: implications for evolution and conservation. Molecular Ecology 15:2141–2152.

Dawson, D. A., M. Åkesson, T. Burke, J. M. Pemberton, J. Slate, and B. Hansson. 2007. Gene order and recombination rate in homologous chromosome regions of the chicken and a passerine bird. Molecular Biology and Evolution 24:1537–1552.

Delmore, K. E., J. W. Fox, and D. E. Irwin. 2012. Dramatic intraspecific differences in migratory routes, stopover sites and wintering areas, revealed using light-level geolocators. Proceedings of the Royal Society of London. Series B: Biological Sciences 279:4582–4589.

Delmore, K. E., and D. E. Irwin. 2014. Hybrid songbirds employ intermediate routes in a migratory divide. Ecology Letters 17:1211–1218.

DePristo, M. A., E. Banks, R. Poplin, K. V. Garimella, J. R. Maguire, C. Hartl, A. A. Philippakis, G. del Angel, M. A. Rivas, M. Hanna, A. McKenna, T. J. Fennell, A. M. Kernytsky, A. Y. Sivachenko, K. Cibulskis, S. B. Gabriel, D. Altshuler, and M. J. Daly. 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nature Genetics 43:491–498.

Dobzhansky, T. 1937a. Genetic nature of species differences. American Naturalist 71: 404–420.

Dobzhansky, T. 1937b. Genetics and the origin of species. Columbia University Press, New York.

Dowling, T. E., and C. L. Secor. 1997. The role of hybridization and introgression in the diversification of animals. Annual Review of Ecology and Systematics 28: 593–619.

Dunning, J. B. 1992. CRC handbook of avian body masses. CRC press, Florida.

Ekstrom, P. A. 2004. An advance in geolocation by light. Memoirs of the National Institute of Polar Research 58:210–226.

Ellegren, H. 2010. Evolutionary stasis: the stable chromosomes of birds. Trends in Ecology & Evolution 25:283–291.

Ellegren, H. 2013. The evolutionary genomics of birds. Annual Review of Ecology, Evolution, and Systematics 44:239–259.

Ellegren, H., L. Smeds, R. Burri, P. I. Olason, N. Backström, T. Kawakami, A. Künstner, H. Mäkinen, K. Nadachowska-Brzyska, A. Qvarnström, S. Uebbing, and J. B. W. Wolf. 2012. The genomic landscape of species divergence in *Ficedula* flycatchers. Nature 491:756–760.

Faaborg, J., R. T. Holmes, A. D. Anders, K. L. Bildstein, K. M. Dugger, S. A. Gauthreaux Jr, P. Heglund, K. A. Hobson, A. E. Jahn, D. H. Johnson, S. C. Latta, D. J. Levey, P. P. Marra, C. L. Merkord, E. Nol, S. I. Rothstein, T. W. Sherry, S. Sillett, F. R. Thompson, and N. Warnock. 2010a. Recent advances in understanding migration systems of New World land birds. Ecological Monographs 80:3–48.

Faaborg, J., R. T. Holmes, A. D. Anders, K. L. Bildstein, K. M. Dugger, S. A. Gauthreaux Jr, P. Heglund, K. A. Hobson, A. E. Jahn, D. H. Johnson, and others. 2010b. Conserving migratory land birds in the New World: Do we know enough? Ecological Applications 20:398–418.

Faircloth, B. C., J. E. McCormack, N. G. Crawford, M. G. Harvey, R. T. Brumfield, and T. C. Glenn. 2012. Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. Systematic Biology 61:717–726.

Feder, J. L., R. Gejji, S. Yeaman, and P. Nosil. 2012. Establishment of new mutations under divergence and genome hitchhiking. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences 367:461–474.

Felger, R. S., and M. Wilson. 1994. Northern Sierra Madre Occidental and its Apachian outliers: a neglected center of biodiversity. Pp. 36–51 *in* F. DeBano, P. Ffolliott, A. Ortega-Rubio, G. Gottfried, and Hamre, eds. Biodiversity and the management of the Madrean Archipelago: The Sky Islands of Southwestern United States and Northwestern Mexico.

Fidler, A. E., K. van Oers, P. J. Drent, S. Kuhn, J. C. Mueller, and B. Kempenaers. 2007. Drd4 gene polymorphisms are associated with personality variation in a passerine bird. Proceedings of the Royal Society of London. Series B: Biological Sciences 274:1685–1691.

Fridolfsson, A. K., and H. Ellegren. 1999. A simple and universal method for molecular sexing of non-ratite birds. Journal of Avian Biology 30:116–121.

Friesen, V. L., A. L. Smith, E. Gómez-Díaz, M. Bolton, R. W. Furness, J. González-Solís, and L. R. Monteiro. 2007. Sympatric speciation by allochrony in a seabird. Proceedings of the National Academy of Sciences 104:18589–18594.

Fudickar, A. M., M. Wikelski, and J. Partecke. 2012. Tracking migratory songbirds: accuracy of light-level loggers (geolocators) in forest habitats. Methods in Ecology and Evolution 3:47–52.

Futschik, A., and C. Schlötterer. 2010. The next generation of molecular markers from massively parallel sequencing of pooled DNA samples. Genetics 186:207–218.

Gagnaire, P. A., V. Albert, B. Jónsson, and L. Bernatchez. 2009. Natural selection influences AFLP intraspecific genetic variability and introgression patterns in Atlantic eels. Molecular Ecology 18:1678–1691.

Gauthreaux Jr, S. A. 1999. Neotropical migrants and the Gulf of Mexico: the view from aloft. Pp. 27–49 *in* K. P. Able, ed. Gatherings of angels: migrating birds and their ecology. Cornell University Press, New York.

Gauthreaux, S. A., C. G. Belser, and C. M. Welch. 2006. Atmospheric trajectories and spring bird migration across the Gulf of Mexico. Journal of Ornithology 147:317–325.

Gay, L., P. A. Crochet, D. A. Bell, and T. Lenormand. 2008. Comparing clines on molecular and phenotypic traits in hybrid zones: a window on tension zone models. Evolution 62:2789–2806.

Gompert, Z., and A. Buerkle. 2011. Bayesian estimation of genomic clines. Molecular Ecology 20:2111–2127.

Gompert, Z., T. L. Parchman, and C. A. Buerkle. 2012. Genomics of isolation in hybrids. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences 367:439–450.

Grant, B. R., and P. R. Grant. 1993. Evolution of Darwin's finches caused by a rare climatic event. Proceedings of the Royal Society of London. Series B: Biological Sciences 251:111–117.

Grant, P. R., and B. R. Grant. 1997. Genetics and the origin of bird species. Proceedings of the National Academy of Sciences 94:7768–7775.

Grant, P. R., and B. R. Grant. 1992. Hybridization of bird species. Science 256:193–197.

Griffin, D. K., L. B. W. Robertson, H. G. Tempest, and B. M. Skinner. 2007. The evolution of the avian genome as revealed by comparative molecular cytogenetics. Cytogenetic and Genome research 117:64–77.

Gwinner, E. 1996. Circadian and circannual programmes in avian migration. Journal of Experimental Biology 199:39–48.

Haavie, J., T. Borge, S. Bures, L. Z. Garamszegi, H. M. Lampe, J. Moreno, A. Qvarnström, J. Török, and G.-P. Sætre. 2004. Flycatcher song in allopatry and sympatry – convergence, divergence and reinforcement. Journal of Evolutionary Biology 17:227–237.

Haldane, J. B. 1922. Sex ratio and unisexual sterility in hybrid animals. Journal of Genetics 12:101–109.

Heckscher, C. M., S. M. Taylor, J. W. Fox, and V. Afanasyev. 2011. Veery (*Catharus fuscescens*) wintering locations, migratory connectivity, and a revision of its winter range using geolocator technology. The Auk 128:531–542.

Hedges, L. V. 1981. Distribution theory for Glass's estimator of effect size and related estimators. Journal of Educational and Behavioral Statistics 6:107–128.

Helbig, A. 1996. Genetic basis, mode of inheritance and evolutionary changes of migratory directions in Palaearctic warblers (Aves: Sylviidae). Journal of Experimental Biology 199:49–55.

Helbig, A. J. 1994. Genetic basis and evolutionary change of migratory direction in a European passerine migrant *Sylvia atricapilla*. Ostrich 65:151–159.

Helbig, A. J. 1991. Inheritance of migratory direction in a bird species: a cross-breeding experiment with SE-and SW-migrating blackcaps (*Sylvia atricapilla*). Behavioral Ecology and Sociobiology 28:9–12.

Hellmann, I., I. Ebersberger, S. E. Ptak, S. Päbo, and M. Przeworski. 2003. A neutral explanation for the correlation of diversity with recombination rates in humans. The American Journal of Human Genetics 72:1527–1535.

Hewitt, G. 2000. The genetic legacy of the Quaternary ice ages. Nature 405:907–913.

Hewitt, G. M. 1988. Hybrid zones-natural laboratories for evolutionary studies. Trends in Ecology & Evolution 3:158–167.

Hill, G. E., and K. McGraw. 2006. Bird coloration. Vol 2. Function and evolution. Harvard University Press, Massachusetts.

Hill, R. 1994. Theory of geolocation by light levels. Pp. 227–236 *in* B.J. Boeuf and R.M. Law, eds. Elephant seals: population ecology, behaviour and physiology. University of California Press, Berkeley.

Hobson, K. A. 2005. Stable isotopes and the determination of avian migratory connectivity and seasonal interactions. The Auk 122:1037–1048.

Hobson, K. A., S. L. Van Wilgenburg, J. Faaborg, J. D. Toms, C. Rengifo, A. L. Sosa, Y. Aubry, and R. Brito Aguilar. 2014. Connecting breeding and wintering grounds of Neotropical migrant songbirds using stable hydrogen isotopes: a call for an isotopic atlas of migratory connectivity. Journal of Field Ornithology 85:237–257.

Hoekstra, P. R. 2010. In search of the elusive behavior gene. Pp. 192–210 *in* B. R. Grant and P. R. Grant, eds. In search of the causes of evolution: From field observations to mechanisms. Princeton University Press, New Jersey.

Hofer, T., M. Foll, and L. Excoffier. 2012. Evolutionary forces shaping genomic islands of population differentiation in humans. BMC Genomics 13:107.

Hostetler, J. A., T. S. Sillett, and P. P. Marra. 2015. Full-annual-cycle population models for migratory birds. The Auk 132:433–449.

Huang, X., and A. Madan. 1999. CAP3: A DNA sequence assembly program. Genome Research 9:868–877.

Hübner, S., E. Rashkovetsky, Y. B. Kim, J. H. Oh, K. Michalak, D. Weiner, A. B. Korol, E. Nevo, and P. Michalak. 2013. Genome differentiation of *Drosophila melanogaster* from a microclimate contrast in Evolution Canyon, Israel. Proceedings of the National Academy of Sciences 110:21059–21064.

Hunt, P. D., B. C. Eliason, and A. Poole. 1999. Blackpoll Warbler (*Dendroica striata*). *in* Birds of North America online. Cornell Lab of Ornithology, New York.

Irwin, D. E., and J. H. Irwin. 2005. Siberian migratory divides: the role of seasonal migration in speciation. Pp. 27–40 *in* R. Greenberg and P. P. Marra, eds. Birds of two worlds: the ecology and evolution of migration. Johns Hopkins University Press, Baltimore.

Irwin, D. E., J. H. Irwin, and T. B. Smith. 2011. Genetic variation and seasonal migratory connectivity in Wilson's warblers (*Wilsonia pusilla*): species-level differences in nuclear DNA between western and eastern populations. Molecular Ecology 20:3102–3115.

Jones, P., D. Binns, H.-Y. Chang, M. Fraser, W. Li, C. McAnulla, H. McWilliam, J. Maslen, A. Mitchell, and G. Nuka. 2014. InterProScan 5: genome-scale protein function classification. Bioinformatics 30:1236–1240.

Jones, S., M. Pfister-Genskow, C. Cirelli, and R. M. Benca. 2008. Changes in brain gene expression during migration in the white-crowned sparrow. Brain Research Bulletin 76:536–544.

Kawakami, T., R. K. Butlin, M. Adams, D. Paull, S. J. B. Cooper, and others. 2008. Genetic analysis of a chromosomal hybrid zone in the Australian morabine grasshoppers (*Vandiemenella*, *viatica* species group). Evolution 63:139–152.

Kelly, J. F., and R. L. Hutto. 2005. An east-west comparison of migration in North American wood warblers. The Condor 107:197–211.

Kloss, B., J. L. Price, L. Saez, J. Blau, A. Rothenfluh, C. S. Wesley, and M. W. Young. 1998. The *Drosophila* clock gene double-time encodes a protein closely related to human casein kinase Iε. Cell 94:97–107.

Kofler, R., P. Orozco-terWengel, N. De Maio, R. V. Pandey, V. Nolte, A. Futschik, C. Kosiol, and C. Schlötterer. 2011a. PoPoolation: a toolbox for population genetic analysis of next generation sequencing data from pooled individuals. PloS ONE 6:e15925.

Kofler, R., R. V. Pandey, and C. Schlötterer. 2011b. PoPoolation2: identifying differentiation between populations using sequencing of pooled DNA samples (Pool-Seq). Bioinformatics 27:3435–3436.

Kofler, R., and C. Schlötterer. 2012. Gowinda: unbiased analysis of gene set enrichment for genome-wide association studies. Bioinformatics 28:2084–2085.

Korf, I. 2004. Gene finding in novel genomes. BMC Bioinformatics 5:59.

Kruuk, L. E. B., S. J. E. Baird, K. S. Gale, and N. H. Barton. 1999. A comparison of multilocus clines maintained by environmental adaptation or by selection against hybrids. Genetics 153:1959–1971.

Kulathinal, R. J., S. M. Bennett, C. L. Fitzpatrick, and M. A. Noor. 2008. Fine-scale mapping of recombination rate in *Drosophila* refines its correlation to diversity and divergence. Proceedings of the National Academy of Sciences 105:10051–10056.

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

Li, H., and R. Durbin. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics 25:1754–1760.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, and R. Durbin. 2009. The sequence alignment/map format and SAMtools. Bioinformatics 25:2078–2079.

Lisovski, S., C. M. Hewson, R. H. Klaassen, F. Korner-Nievergelt, M. W. Kristensen, and S. Hahn. 2012. Geolocation by light: accuracy and precision affected by environmental factors. Methods in Ecology and Evolution 3:603–612.

Lovette, I. J. 2005. Glacial cycles and the tempo of avian speciation. Trends in Ecology & Evolution 20:57–59.

Lovette, I. J., S. M. Clegg, and T. B. Smith. 2004. Limited utility of mtDNA markers for determining connectivity among breeding and overwintering locations in three Neotropical migrant birds. Conservation Biology 18:156–166.

Lovette, I. J., J. L. Pérez-Emán, J. P. Sullivan, R. C. Banks, I. Fiorentino, S. Córdoba-Córdoba, M. Echeverry-Galvis, F. K. Barker, K. J. Burns, J. Klicka, and others. 2010. A comprehensive multilocus phylogeny for the wood-warblers and a revised classification of the Parulidae (Aves). Molecular Phylogenetics and Evolution 57:753–770.

Lowry, D. B., J. L. Modliszewski, K. M. Wright, C. A. Wu, and J. H. Willis. 2008. The strength and genetic basis of reproductive isolating barriers in flowering plants. Philosophical Transactions of the Royal Society B: Biological Sciences 363:3009–3021.

Lundberg, M., S. Åkesson, and S. Bensch. 2011. Characterization of a divergent chromosome region in the willow warbler *Phylloscopus trochilus* using avian genomic resources. Journal of Evolutionary Biology 24:1241–1253.

Lundberg, M., J. Boss, B. Canbäck, M. Liedvogel, K. W. Larson, M. Grahn, S. Åkesson, S. Bensch, and A. Wright. 2013. Characterisation of a transcriptome to find sequence differences between two differentially migrating subspecies of the willow warbler *Phylloscopus trochilus*. BMC Genomics 14:330.

Lunter, G., and M. Goodson. 2011. Stampy: a statistical algorithm for sensitive and fast mapping of Illumina sequence reads. Genome Research 21:936–939.

Mack, D., W. Yong, and A. Poole. 2000. Swainson's Thrush (*Catharus ustulatus*). *in* Birds of North America Online. Cornell Lab of Ornithology, New York.

Maia, R., C. M. Eliason, P.-P. Bitton, S. M. Doucet, and M. D. Shawkey. 2013. pavo: an R package for the analysis, visualization and organization of spectral data. Methods in Ecology and Evolution 4:906–913.

Mallet, J. 2007. Hybrid speciation. Nature 446:279–283.

Mank, J. E., E. Axelsson, and H. Ellegren. 2007. Fast-X on the Z: rapid evolution of sex-linked genes in birds. Genome Research 17:618–624.

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

Martin, P. R., R. Montgomerie, and S. C. Lougheed. 2010. Rapid sympatry explains greater color pattern divergence in high latitude birds. Evolution 64:336–347.

Maynard Smith, J. M., and J. Haigh. 1974. The hitch-hiking effect of a favourable gene. Genetical Research 23:23–35.

Mayr, E. 1942. Systematics and the origin of species, from the viewpoint of a zoologist. Columbia University Press, New York.

McCarthy, E. M. 2006. Handbook of avian hybrids of the world. Oxford University Press, New York.

McDevitt, A. D., S. Mariani, M. Hebblewhite, N. J. Decesare, L. Morgantini, D. Seip, B. V. Weckworth, and M. Musiani. 2009. Survival in the Rockies of an endangered hybrid swarm from diverged caribou (*Rangifer tarandus*) lineages. Molecular Ecology 18:665–679.

McKinnon, E. A., K. C. Fraser, and B. J. Stutchbury. 2013. New discoveries in landbird migration using geolocators, and a flight plan for the future. The Auk 130:211–222.

McRae, B. H., B. G. Dickson, T. H. Keitt, and V. B. Shah. 2008. Using circuit theory to model connectivity in ecology, evolution, and conservation. Ecology 89:2712–2724.

Mettler, R., H. M. Schaefer, N. Chernetsov, W. Fiedler, K. A. Hobson, M. Ilieva, E. Imhof, A. Johnsen, S. C. Renner, and G. Rolshausen. 2013. Contrasting patterns of genetic differentiation among blackcaps (*Sylvia atricapilla*) with divergent migratory orientations in Europe. PloS ONE 8:e81365.

Metzker, M. L. 2010. Sequencing technologies—the next generation. Nature Reviews Genetics 11:31–46.

Møller, A. P., L. Z. Garamszegi, J. M. Peralta-Sánchez, and J. J. Soler. 2011. Migratory divides and their consequences for dispersal, population size and parasite–host interactions. Journal of Evolutionary Biology 24:1744–1755.

Montgomerie, R. 2008. CLR, Version 1.03. Queen's University, Kingston, Ontario.

Moore, F. R. 1999. Cheniers of Louisiana and the stopover ecology of migrant landbirds. Pp. 51–62 *in* K. Able, ed. Gatherings of angels: migrating birds and their ecology. Cornell University Press, New York.

Mueller, J. C., F. Pulido, and B. Kempenaers. 2011. Identification of a gene associated with avian migratory behaviour. Proceedings of the Royal Society B: Biological Sciences 278:2848–2856.

Nadeau, N. J., S. H. Martin, K. M. Kozak, C. Salazar, K. K. Dasmahapatra, J. W. Davey, S. W. Baxter, M. L. Blaxter, J. Mallet, and C. D. Jiggins. 2013. Genome-wide patterns of divergence and gene flow across a butterfly radiation. Molecular Ecology 22:814–826.

Nei, M. 1987. Molecular evolutionary genetics. Columbia University Press, New York.

Nei, M., and S. Kumar. 2000. Molecular evolution and phylogenetics. Oxford University Press, West Sussex.

Newton, I. 2008. The migration ecology of birds. Academic Press, London.

Noor, M. A., and S. M. Bennett. 2009. Islands of speciation or mirages in the desert? Examining the role of restricted recombination in maintaining species. Heredity 103:439–444.

Noor, M. A. F. 1999. Reinforcement and other consequences of sympatry. Heredity 83:503–508.

Noor, M. A. F., and J. L. Feder. 2006. Speciation genetics: evolving approaches. Nature Reviews Genetics 7:851–861.

Norris, D. R., P. P. Marra, R. Montgomerie, T. K. Kyser, and L. M. Ratcliffe. 2004. Reproductive effort, molting latitude, and feather color in a migratory songbird. Science 306:2249–2250.

Nosil, P. 2012. Ecological speciation. Oxford University Press, New York.

Nosil, P., and J. L. Feder. 2012. Genomic divergence during speciation: causes and consequences. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences 367:332–342.

Nosil, P., D. J. Funk, and D. Ortiz-Barrientos. 2009. Divergent selection and heterogeneous genomic divergence. Molecular Ecology 18:375–402.

O'Corry-Crowe, G. M., R. S. Suydam, A. Rosenberg, K. J. Frost, and A. E. Dizon. 1997. Phylogeography, population structure and dispersal patterns of the beluga whale *Delphinapterus leucas* in the western Nearctic revealed by mitochondrial DNA. Molecular Ecology 6:955–970.

Orr, H. A. 2005. The genetic basis of reproductive isolation: Insights from *Drosophila*. Proceedings of the National Academy of Sciences 102:6522–6526.

Parchman, T. L., Z. Gompert, M. J. Braun, R. T. Brumfield, D. B. McDonald, J. A. C. Uy, G. Zhang, E. D. Jarvis, B. A. Schlinger, and C. A. Buerkle. 2013. The genomic consequences of adaptive divergence and reproductive isolation between species of manakins. Molecular Ecology 22:3304–3317.

Parra, G., P. Agarwal, J. F. Abril, T. Wiehe, J. W. Fickett, and R. Guigó. 2003. Comparative gene prediction in human and mouse. Genome Research 13:108–117.

Pérez-Tris, J., S. Bensch, R. Carbonell, A. Helbig, and J. L. Tellería. 2004. Historical diversification of migration patterns in a passerine bird. Evolution 58:1819–1832.

Pfennig, D. W., and K. S. Pfennig. 2010. Character displacement and the origins of diversity. American Naturalist 176:S26–S44.

Pfennig, K. S. 1998. The evolution of mate choice and the potential for conflict between species and mate–quality recognition. Proceedings of the Royal Society B: Biological Sciences 265:1743–1748.

Phillips, A. R. 1975. The migrations of Allen's and other hummingbirds. The Condor 77:196–205.

Presgraves, D. C. 2008. Sex chromosomes and speciation in *Drosophila*. Trends in Genetics 24:336–343.

Price, T. 2008. Speciation in birds. Roberts and Company Publishers, Colorado.

Price, T. D., and M. M. Bouvier. 2002. The evolution of F1 postzygotic incompatibilities in birds. Evolution 56:2083–2089.

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

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

Pyle, P., S. N. G. Howell, and S. Ruck. 1997. Identification Guide to North American Birds: A Compendium of Information on Identifying, Ageing, and Sexing Waterbirds, Diurnal Raptors, and Gallinaceous Birds in the Hand. Anatidae to Alcidae. Slate Creek Press.

Rabosky, D. L., and D. R. Matute. 2013. Macroevolutionary speciation rates are decoupled from the evolution of intrinsic reproductive isolation in *Drosophila* and birds. Proceedings of the National Academy of Sciences 110:15354–15359.

Rappole, J. 1995. The ecology of migrant birds. Smithsonian Institution Press, Washington.

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.

R Development Core Team. 2013, 2014. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna.

Renaut, S., C. J. Grassa, S. Yeaman, B. T. Moyers, Z. Lai, N. C. Kane, J. E. Bowers, J. M. Burke, and L. H. Rieseberg. 2013. Genomic islands of divergence are not affected by geography of speciation in sunflowers. Nature Communications 4:1827.

Reudink, M. W., P. P. Marra, P. T. Boag, and L. M. Ratcliffe. 2009. Plumage coloration predicts paternity and polygyny in the American redstart. Animal Behaviour 77:495–501.

Ridgely, R., T. Allnutt, T. Brooks, D. McNicol, D. Mehlman, B. Young, J. Zook, and BirdLife International. 2011. Digital distribution maps of the birds of the western hemisphere, version 4.0. *in* BirdLife International and NatureServe Bird species distribution maps of the world. BirdLife International and NatureServe, Cambridge, UK and Arlington, USA.

Rieseberg, L., and A. Buerkle. 2002. Genetic mapping in hybrid zones. The American Naturalist 159:S36–S50.

Roesti, M., D. Moser, and D. Berner. 2013. Recombination in the threespine stickleback genome-patterns and consequences. Molecular Ecology 22:3014–3027.

Rohwer, S., L. Butler, and D. Froehlich. 2005. Ecology and demography of east-west differences in molt scheduling of Neotropical migrant passerines. Pp. 87–105 *in* R. Greenburg and P. P. Marra, eds. Birds of two worlds: the ecology and evolution of migration. Johns Hopkins University Press, Baltimore.

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

Rolshausen, G., G. Segelbacher, K. A. Hobson, and H. M. Schaefer. 2009. Contemporary evolution of reproductive isolation and phenotypic divergence in sympatry along a migratory divide. Current Biology 19:2097–2101.

Ruegg, K. 2008. Genetic, morphological, and ecological characterization of a hybrid zone that spans a migratory divide. Evolution 62:452–466.

Ruegg, K., E. C. Anderson, J. Boone, J. Pouls, and T. B. Smith. 2014. A role for migration-linked genes and genomic islands in divergence of a songbird. Molecular Ecology 23:4757–4769.

Ruegg, K. C., R. J. Hijmans, and C. Moritz. 2006a. Climate change and the origin of migratory pathways in the Swainson's thrush, *Catharus ustulatus*. Journal of Biogeography 33:1172–1182.

Ruegg, K. C., and T. B. Smith. 2002. Not as the crow flies: a historical explanation for circuitous migration in Swainson's thrush (*Catharus ustulatus*). Proceedings of the Royal Society of London. Series B: Biological Sciences 269:1375–1381.

Ruegg, K., H. Slabbekoorn, S. Clegg, and T. Smith. 2006b. Divergence in mating signals correlates with ecological variation in the migratory songbird, Swainson's thrush (*Catharus ustulatus*). Molecular Ecology 15:3147–3156.

Rundle, H. D., and P. Nosil. 2005. Ecological speciation. Ecology Letters 8:336–352.

Saetre, G.-P., T. Borge, K. Lindroos, J. Haavie, B. C. Sheldon, C. Primmer, and A.-C. Syvänen. 2003. Sex chromosome evolution and speciation in *Ficedula* flycatchers. Proceedings of the Royal Society of London. Series B: Biological Sciences 270:53–59.

Safran, R., S. Flaxman, M. Kopp, D. E. Irwin, D. Briggs, M. R. Evans, W. C. Funk, D. A. Gray, E. A. Hebets, and N. Seddon. 2012. A robust new metric of phenotypic distance to estimate and compare multiple trait differences among populations. Current Zoology 58:423–436.

Schielzeth, H., W. Forstmeier, B. Kempenaers, and H. Ellegren. 2012. QTL linkage mapping of wing length in zebra finch using genome-wide single nucleotide polymorphisms markers. Molecular Ecology 21:329–339.

Schluter, D. 2001. Ecology and the origin of species. Trends in Ecology & Evolution 16:372–380.

Schluter, D. 2009. Evidence for ecological speciation and its alternative. Science 323:737–741.

Schluter, D. 2000. The ecology of adaptive radiation. Oxford University Press, New York.

Schluter, D., and G. L. Conte. 2009. Genetics and ecological speciation. Proceedings of the National Academy of Sciences 106:9955–9962.

Schwenk, K., N. Brede, and B. Streit. 2008. Introduction. Extent, processes and evolutionary impact of interspecific hybridization in animals. Philosophical Transactions of the Royal Society B: Biological Sciences 363:2805–2811.

Seddon, N., C. A. Botero, J. A. Tobias, P. O. Dunn, H. E. A. MacGregor, D. R. Rubenstein, J. A. C. Uy, J. T. Weir, L. A. Whittingham, and R. J. Safran. 2013. Sexual selection accelerates signal evolution during speciation in birds. Proceedings of the Royal Society of London. Series B: Biological Sciences 280:1065.

Seehausen, O., R. K. Butlin, I. Keller, C. E. Wagner, J. W. Boughman, P. A. Hohenlohe, C. L. Peichel, G.-P. Saetre, C. Bank, Åke Brännström. 2014. Genomics and the origin of species. Nature Reviews Genetics 15:176–192.

Shields, D. C., P. M. Sharp, D. G. Higgins, and F. Wright. 1988. "Silent" sites in *Drosophila* genes are not neutral: evidence of selection among synonymous codons. Molecular Biology of Evolution 5:704–716.

Skagen, S. K., J. F. Kelly, C. van Riper III, R. L. Hutto, D. M. Finch, D. J. Krueper, and C. P. Melcher. 2005. Geography of spring landbird migration through riparian habitats in southwestern North America. The Condor 107:212–227.

Slabbekoorn, H., and T. B. Smith. 2002. Bird song, ecology and speciation. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences 357:493–503.

Smit, A., R. Hubley, and P. Green. 1996. RepeatMasker Open-3.0. <http://www.repeatmasker.org>.

Smith, J. W., and C. W. Benkman. 2007. A coevolutionary arms race causes ecological speciation in Crossbills. The American Naturalist 169:455–465.

Smith, T. B., S. M. Clegg, M. Kimura, K. Ruegg, B. Mila, and I. Lovette. 2005. Molecular genetic approaches to linking breeding and overwintering areas in five neotropical migrant passerines. Pp. 222–234 *in* R. Greenburg and P. Marra, eds. Birds of two worlds: the ecology and evolution of migration. Johns Hopkins University Press, Baltimore.

Soria-Carrasco, V., Z. Gompert, A. A. Comeault, T. E. Farkas, T. L. Parchman, J. S. Johnston, C. A. Buerkle, J. L. Feder, J. Bast, T. Schwander, S. P. Egan, B. J. Crespi, and P. Nosil. 2014. Stick insect genomes reveal natural selection's role in parallel speciation. Science 344:738–742.

Sousa, V., and J. Hey. 2013. Understanding the origin of species with genome-scale data: modelling gene flow. Nature Reviews Genetics 14:404–414.

Stanke, M., A. Tzvetkova, and B. Morgenstern. 2006. AUGUSTUS at EGASP: using EST, protein and genomic alignments for improved gene prediction in the human genome. Genome Biology 7:S11.

Stanley, C. Q., M. MacPherson, K. C. Fraser, E. A. McKinnon, and B. J. M. Stutchbury. 2012. Repeat tracking of individual songbirds reveals consistent migration timing but flexibility in route. PLoS ONE 7:e40688.

Stapley, J., T. R. Birkhead, T. Burke, and J. Slate. 2008. A linkage map of the zebra finch *Taeniopygia guttata* provides new insights into avian genome evolution. Genetics 179:651–667.

Stapley, J., T. R. Birkhead, T. Burke, and J. Slate. 2010a. Pronounced inter- and intrachromosomal variation in linkage disequilibrium across the zebra finch genome. Genome Resources 20:496–502.

Stapley, J., J. Reger, P. G. D. Feulner, C. Smadja, J. Galindo, R. Ekblom, C. Bennison, A. D. Ball, A. P. Beckerman, and J. Slate. 2010b. Adaptation genomics: the next generation. Trends in Ecology & Evolution 25:705–712.

Steinmeyer, C., J. C. Mueller, and B. Kempenaers. 2009. Search for informative polymorphisms in candidate genes: clock genes and circadian behaviour in blue tits. Genetica 136:109–117.

Stoddard, M. C., and R. O. Prum. 2008. Evolution of avian plumage color in a tetrahedral color space: a phylogenetic analysis of new world buntings. The American Naturalist 171:755–776.

Storchová, R., J. Reif, and M. W. Nachman. 2010. Female heterogamety and speciation: reduced introgression of the Z chromosome between two species of nightingales. Evolution 64:456–471.

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 323:896–896.

Swenson, N. G., and D. J. Howard. 2004. Do suture zones exist? Evolution 58:2391–2397.

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

Tajima, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123:585–595.

Templeton, A. R. 1981. Mechanisms of speciation–a population genetic approach. Annual Review of Ecology and Systematics 12: 23–48.

Tobias, J. A., N. Seddon, C. N. Spottiswoode, J. D. Pilgrim, L. D. Fishpool, and N. J. Collar. 2010. Quantitative criteria for species delimitation. Ibis 152:724–746.

Tøttrup, A. P., R. H. G. Klaassen, R. Strandberg, K. Thorup, M. W. Kristensen, P. S. Jørgensen, J. Fox, V. Afanasyev, C. Rahbek, and T. Alerstam. 2012. The annual cycle of a trans-equatorial Eurasian–African passerine migrant: different spatio-temporal strategies for autumn and spring migration. Proceedings of the Royal Society B: Biological Sciences 279:1008–1016.

Turner, T. L., and M. W. Hahn. 2010. Genomic islands of speciation or genomic islands and speciation? Molecular Ecology 19:848–850.

Uy, J. A. C., R. G. Moyle, and C. E. Filardi. 2009. Plumage and song differences mediate species recognition between incipient flycatcher species of the Solomon Islands. Evolution 63:153–164.

Veen, T., N. Svedin, J. T. Forsman, M. B. Hjernquist, A. Qvarnström, K. A. T. Hjernquist, J. Träff, and M. Klaassen. 2007. Does migration of hybrids contribute to post-zygotic isolation in flycatchers? Proceedings of the Royal Society B: Biological Sciences 274:707–712.

Via, S. 2012. Divergence hitchhiking and the spread of genomic isolation during ecological speciation-with-gene-flow. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences 367:451–460.

Via, S. 2009. Natural selection in action during speciation. Proceedings of the National Academy of Sciences 106:9939–9946.

Wager-Smith, K., and S. A. Kay. 2000. Circadian rhythm genetics: from flies to mice to humans. Nature Genetics 26:23–27.

Wahlberg, P., L. Strömstedt, X. Tordoir, M. Foglio, S. Heath, D. Lechner, A. R. Hellström, M. Tixier-Boichard, M. Lathrop, and I. G. Gut. 2007. A high-resolution linkage map for the Z chromosome in chicken reveals hot spots for recombination. Cytogenetic and Genome Research 117:22–29.

Warren, W. C., D. F. Clayton, H. Ellegren, A. P. Arnold, L. W. Hillier, A. Künstner, S. Searle, S. White, A. J. Vilella, and S. Fairley. 2010. The genome of a songbird. Nature 464:757–762.

Webster, M. S., and P. P. Marra. 2005. The importance of understanding migratory connectivity and seasonal interactions. Pp. 199–209 *in* R. Greenberg and P. P. Marra, eds. Birds of two worlds: the ecology and evolution of ,igration. Johns Hopkins University Press, Baltimore.

Webster, M. S., P. P. Marra, S. M. Haig, S. Bensch, and R. T. Holmes. 2002. Links between worlds: unraveling migratory connectivity. Trends in Ecology & Evolution 17:76–83.

Weir, J. T., and D. Schluter. 2004. Ice sheets promote speciation in boreal birds. Proceedings of the Royal Society of London. Series B: Biological Sciences 271:1881–1887.

Weir, J. T., and D. Wheatcroft. 2011. A latitudinal gradient in rates of evolution of avian syllable diversity and song length. Proceedings of the Royal Society B: Biological Sciences 278:1713–1720.

Weir, J. T., D. J. Wheatcroft, and T. D. Price. 2012. The role of ecological constraint in driving the evolution of avian song frequency across a latitudinal gradient. Evolution 66:2773–2783.

Wiltschko, R., and W. Wiltschko. 1999. Compass orientation as a basic element in avian orientation and navigation. Pp. 259–293 *in* Golledge, R.G., ed. Wayfinding behavior. Johns Hopkins University Press, Baltimore.

Wolf, J. B., J. Lindell, and N. Backström. 2010. Speciation genetics: current status and evolving approaches. Philosophical Transactions of the Royal Society B: Biological Sciences 365:1717–1733.

Wu, C.-I. 2001. Genes and speciation. Journal of Evolutionary Biology 14:889–891.

Wu, C.-I., and C.-T. Ting. 2004. Genes and speciation. Nature Reviews Genetics 5:114–122.

Appendices

Appendix A Supplementary material for Chapter 2

A.1 Additional details regarding the analysis of data for two birds tracked from pure populations.

One of the returned devices moved down the back of the bird during deployment (shown in green on Figure 2.1). We examined light transitions recorded by this device and inferred that this event most likely occurred in August, as light transitions were non-linear throughout the day for this month. To compensate for this change, we split the dataset in two and used two sun elevation angles to estimate daily longitude and latitude. We used the sun elevation calculated on the breeding grounds in May for the first dataset and the sun elevation calculated in July of the following year for the second dataset. This bird remained on the breeding grounds until the beginning of September, allowing us to evaluate how accurate this treatment of the data was; we calculated a mean position for data before and after correction and observed similar points, N48.285 W121.98 and N49.117 W122.

One of the returned devices experienced a large amount of clock drift (24 minutes, shown in red on Figure 2.1). By stepping through the data we were able to identify the specific day where a great deal of this drift appears to have occurred; after July 10, 2010 there was a large shift west in our estimates of longitude. We calculated a mean position for the data before and after this shift and observed that there was a difference of 3 degrees longitude between these points. Three degrees longitude is equivalent to 12 minutes of time. Accordingly, we removed 12 minutes from all transitions after July 10. We calculated a mean position for data before and after correction and observed similar points, 51.811 W118.905 and N 51.431 W119.104.

A.2 Daily estimates of longitude and latitude for Swainson's thrushes tracked from pure populations. Standard deviations (SD) are shown and were not calculated when data from only one day was available. Latitude was not estimated for at least thirty days around the equinoxes; estimates from these dates are based only on longitude. Colours correspond to those used in Figure 2.1, with data for thrushes from the inland subspecies group shown in warm colours (red, orange, yellow, green) and from the coastal subspecies group shown in cool colours (blue, purple, pink, turquoise, maroon).

Dates	Longitude (^{o}W)	SD	Latitude (^{o}N)	SD
Maroon				
Jun-12 to Jul-15	122.21	1.26	49.37	0.80
Jul-17 to Sep-22	121.41	1.33	47.85	1.86
Sep-23	113.30			
Sep-24	109.39			
Sep-26 to Oct-03	105.40	1.02		
Oct-04	99.83			
Oct-07 to Oct-16	96.04	0.61	17.95	2.11
Oct-17 to Dec-28	92.30	0.74	18.13	2.05
Dec-29 to Apr-16	90.50	0.71	16.96	2.62
Apr-17	96.97		16.16	
Apr-18 to May-03	99.44	0.47	17.99	1.55
May-04	102.55		24.34	
May-05	103.95		27.35	
May-06	107.59		27.24	
May-07	109.61		32.28	
May-08	110.87		32.38	
May-09	112.51		33.67	
May-10 to May-11	113.03	0.18	35.14	0.27
May-12 to May-14	115.20	0.63	36.40	0.90
May-15 to May-18	117.87	0.64	36.95	0.98
May-20	122.36		40.26	
May-21 to May-31	122.52	0.79	44.82	0.95
Jun-01 to Jun-03	123.50	0.30	48.47	0.60

Dates	Longitude (^{o}W)	SD	Latitude (^{o}N)	SD
Turquoise				
Jun-09 to Jul-25	122.30	0.93	50.17	0.73
Jul-28 to Sep-20	122.45	0.99	47.03	1.62
Sep-21	117.75			
Sep-22 to Oct-09	103.49	4.37		
Oct-18 to Dec-21	84.11	1.88	15.37	4.60
Dec-22 to Apr-06	89.57	1.08	16.53	2.48
Apr-10 to May-03	98.43	1.27	16.45	2.29
May-04	102.30		23.04	
May-05	104.32		24.59	
May-06	107.59		27.82	
May-07 to May-11	109.73	0.45	30.36	0.83
May-12	112.66		34.13	
May-13	116.04		34.86	
May-14 to May-20	120.91	1.15	37.86	1.21
May-21 to May-27	121.24	0.92	44.18	1.27
May-28 to Jun-08	123.98	1.73	50.42	1.85
Blue				
Jun-12 to Jul-28	122.71	0.82	49.45	0.53
Jul-29 to Sep-22	121.15	1.10	47.16	0.82
Sep-23	111.80			
Sep-24 to Oct-02	106.24	1.00		
Oct-03	102.00			
Oct-04	98.58			
Oct-05 to Oct-13	96.10	0.47	16.34	2.53
Oct-14 to Apr-17	90.35	1.12	19.04	2.87
Apr-18	93.03		18.43	
Apr-19	97.84		16.71	
Apr-21 to May-07	104.09	0.47	19.03	3.11
May-08	106.87		26.49	
May-09	108.26		31.23	
May-10	112.15		32.57	
May-11	114.03		33.83	
May-12	117.66		36.28	
May-13	117.91		39.81	
May-14 to May-19	122.10	1.65	40.32	0.74
May-20	121.86		43.62	
May-21	123.72		44.84	

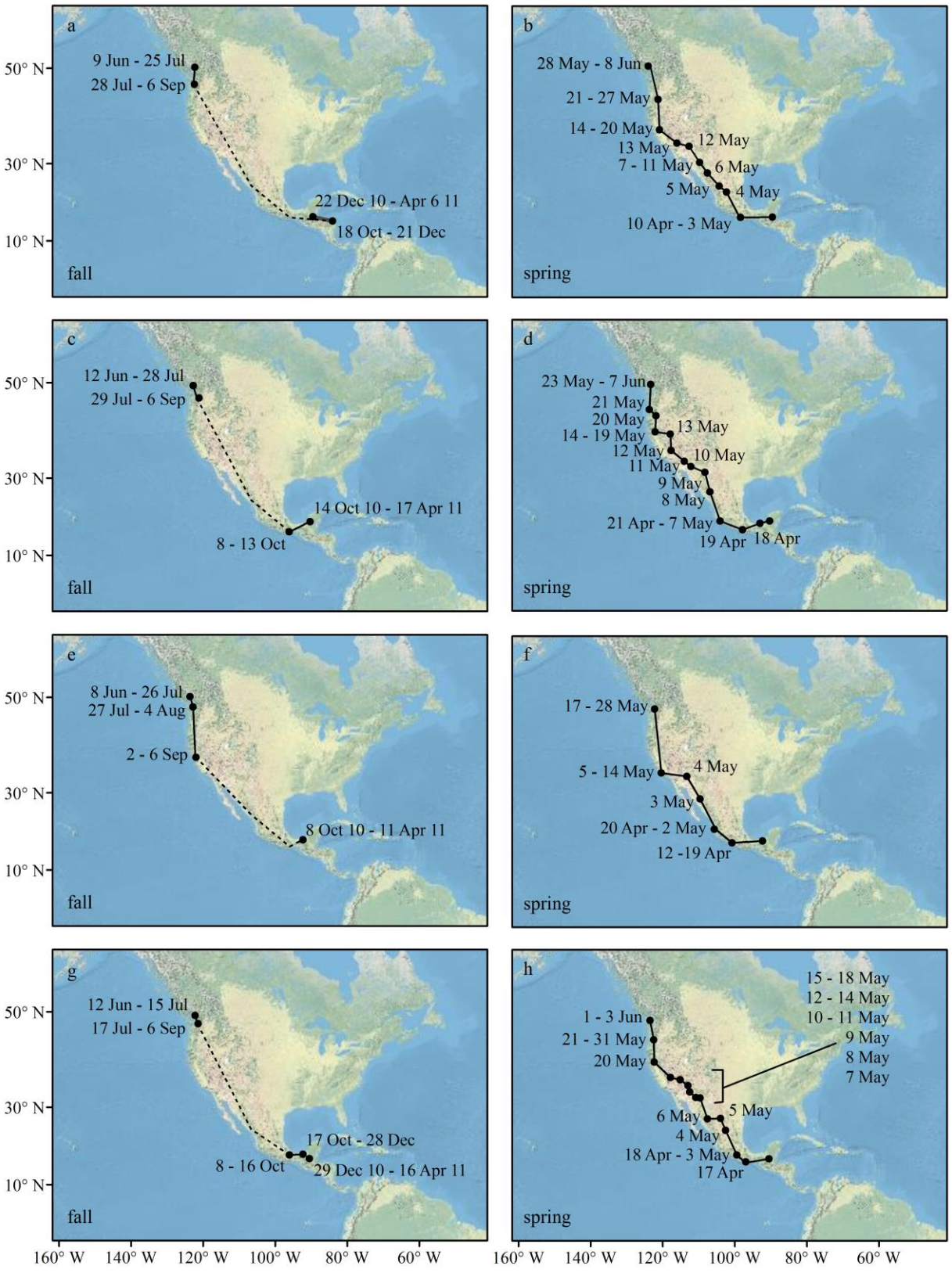
Dates	Longitude (^{o}W)	SD	Latitude (^{o}N)	SD
May-23 to Jun-07	123.35	1.06	49.57	0.96
Pink				
Jun-12 to Jul-24	122.06	1.01	49.32	0.89
Jul-26 to Sep-01	121.27	1.34	46.91	2.89
Sep-02 to Sep-28	121.63	1.17	42.84	0.58
Sep-29 to Oct-08	108.70	3.00	29.04	
Oct-09 to Oct-17	99.51	0.49	20.93	7.80
Oct-18	94.09		18.28	
Oct-19	90.38		15.11	
Oct-20 to Dec-23	86.50	0.86	15.83	4.13
Dec-24 to Apr-01	90.61	1.08	18.80	7.81
Apr-03 to May-01	99.24	0.96	17.06	8.71
May-02	102.62		24.82	
May-03	106.28		28.49	
May-04	111.30		33.07	
May-05	115.07		31.03	
May-06 to May-11	120.63	1.12	36.64	3.73
May-12	123.91		42.94	
May-13	121.91		45.05	
Purple				
Jun-08 to Jul-26	123.61	1.26	50.10	0.78
Jul-27 to Aug-04	122.83	0.65	48.25	0.72
Sep-02 to Sep-14	121.98	1.46	38.11	1.00
Sep-15	118.22			
Sep-16 to Oct-01	101.93	3.78		
Oct-02 to Apr-11	92.33	1.22	18.13	6.15
Apr-12 to Apr-19	100.76	1.80	17.64	8.22
Apr-20 to May-02	105.71	0.70	21.21	3.39
May-03	109.65		28.78	
May-04	113.30		34.10	
May-05 to May-14	120.42	2.45	34.94	1.51
May-17 to May-28	122.27	1.03	48.10	1.26
Orange				
Jun-15 to Jul-31	121.35	0.82	52.54	0.56
Aug-01 to Sep-09	120.48	0.98	49.03	1.19
Sep-12 to Sep-24	116.92	1.67		

Dates	Longitude (^{o}W)	SD	Latitude (^{o}N)	SD
Sep-25	105.72			
Sep-26	99.93			
Sep-28 to Oct-13	88.70	0.82	21.30	2.33
Oct-14	84.75			
Oct-15	80.06			
Oct-16	77.36		8.16	
Oct-17 to Apr-07	72.74	0.65	8.63	2.50
Apr-08	74.77			
Apr-09	77.34			
Apr-10 to Apr-17	82.67	0.79	8.63	6.58
Apr-18 to Apr-27	88.62	1.28	13.67	8.71
Apr-28	91.87		16.32	
Apr-30 to May-03	97.76	1.55	25.61	2.14
May-04 to May-16	95.42	1.46	35.04	1.42
May-17	103.40		45.12	
May-18	109.64		47.66	
May-19	114.00		46.48	
May-20	117.36		48.00	
May-21	120.72		49.38	
May-22 to Jun-13	121.05	0.89	52.12	0.40
Yellow				
Jun-15 to Aug-28	120.46	0.64	52.04	0.60
Aug-30 to Sep-04	118.95	0.60	50.78	0.32
Sep-05 to Sep-08	113.56	0.74	49.38	0.82
Sep-11	95.73			
Sep-12 to Oct-06	85.18	1.86		
Oct-07	79.30			
Oct-08 to Apr-10	72.42	1.22	12.70	4.52
Apr-11 to Apr-17	81.32	2.00	8.48	1.97
Apr-18 to Apr-30	90.53	1.34	14.13	5.20
May-01 to May-14	96.18	2.00	32.73	2.38
May-15	103.78		39.29	
May-16	104.28		44.46	
May-17	111.27		45.80	
May-18	116.26		46.53	
May-19 to May-20	118.19	0.08	49.21	0.34
May-21 to Jun-12	120.43	0.84	51.93	0.42

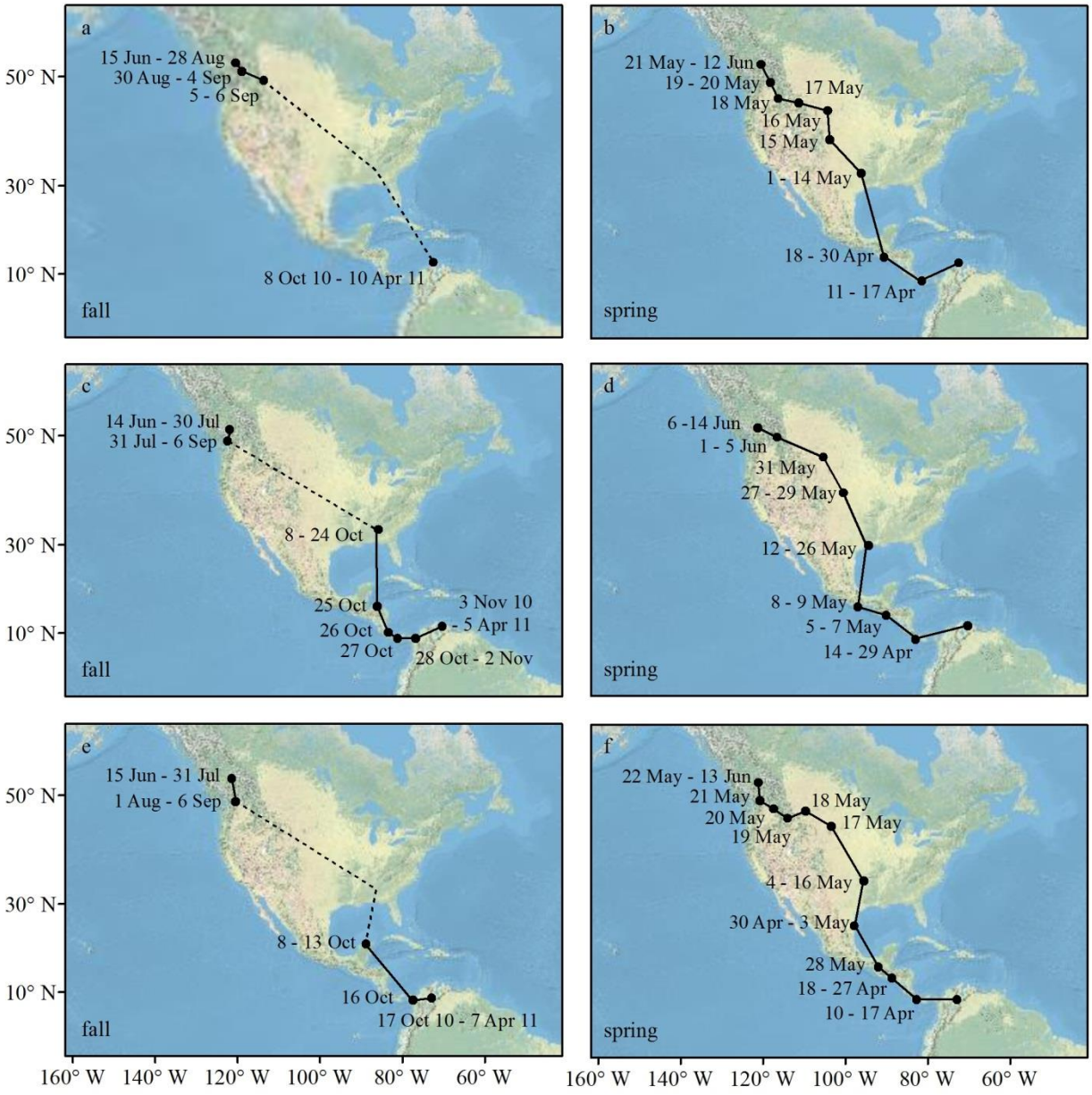
Dates	Longitude (^{o}W)	SD	Latitude (^{o}N)	SD
Red				
Jun-16 to Jul-31	119.01	0.79	51.62	0.53
Aug-01 to Sep-14	119.69	0.66	50.10	1.17
Sep-17 to Sep-24	113.98	1.41		
Sep-25	104.60			
Sep-26	100.18			
Sep-28	92.72			
Sep-30 to Oct-13	89.03	1.74	41.19	0.47
Oct-14 to Oct-24	87.20	0.79	32.59	3.56
Oct-26	83.63		15.37	
Oct-31 to Nov-24	80.36	0.74	8.80	3.16
Nov-26	77.41		8.32	
Nov-27 to Mar-26	71.97	1.75	12.26	6.18
Mar-28 to Apr-11	81.70	0.64	8.00	4.98
Apr-14 to Apr-30	88.56	1.22	13.67	10.02
May-01	93.84		16.30	
May-02	94.25			
May-03 to May-04	97.92	0.19	21.87	0.95
May-05 to May-10	98.38	0.52	26.83	0.38
May-11	99.03		30.00	
May-12 to May-20	97.46	1.50	33.43	0.92
May-21 to May-31	98.92	2.54	39.26	1.80
Jun-02 to Jun-05	109.40	1.66	43.94	1.81
Jun-06	113.20		49.30	
Jun-09 to Jun-12	117.27	1.61	51.01	1.33
Green				
Jun-14 to Jul-30	121.82	1.11	50.95	0.53
Jul-31 to Sep-14	122.32	1.18	49.12	1.58
Sep-15	112.09			
Sep-16	100.80			
Sep-17 to Sep-24	93.34	2.02		
Sep-26 to Oct-24	85.81	2.42	33.07	4.15
Oct-25	86.10		16.19	
Oct-26	83.38		10.06	
Oct-27	81.03		8.64	
Oct-28 to Nov-02	76.73	0.90	8.64	3.16
Nov-03 to Apr-04	70.20	1.34	11.48	3.27
Apr-14 to Apr-29	82.87	2.05	8.32	12.05

Dates	Longitude (^{o}W)	SD	Latitude (^{o}N)	SD
May-05 to May-07	89.97	0.84	13.98	3.69
May-08 to May-09	96.94		15.99	7.15
May-12 to May-26	94.35	2.44	29.74	4.65
May-27 May-29	100.51	1.25	40.12	2.62
May-31	105.33		46.43	
Jun-01 to Jun-05	116.46	1.60	49.67	3.97
Jun-06 to Jun-14	121.23	0.85	51.14	0.56

A.3 Annual cycle of the remaining coastal Swainson's thrushes tracked from pure populations. Dates, stopover sites, breeding and wintering locations are shown for fall (left panels) and spring migration (right panels). Missing dates indicate periods for which the location of the bird could not be estimated. Dashed lines link locations where latitude could not be estimated around the equinox periods. Thick gray lines link locations on the wintering grounds. Panels a and b correspond with routes shown in turquoise on Figure 2.1, c and d with routes shown in blue, e and f with routes shown in purple and g and h with routes shown in maroon.


a
9 Jun - 25 Jul
28 Jul - 6 Sep
22 Dec 10 - Apr 6 11
18 Oct - 21 Dec
fall
b
28 May - 8 Jun
21 - 27 May
14 - 20 May
13 May
12 May
7 - 11 May
6 May
5 May
4 May
10 Apr - 3 May
spring
c
12 Jun - 28 Jul
29 Jul - 6 Sep
14 Oct 10 - 17 Apr 11
8 - 13 Oct
fall
d
23 May - 7 Jun
21 May
20 May
14 - 19 May
13 May
12 May
10 May
11 May
9 May
8 May
21 Apr - 7 May
19 Apr
18 Apr
spring
e
8 Jun - 26 Jul
27 Jul - 4 Aug
2 - 6 Sep
8 Oct 10 - 11 Apr 11
fall
f
17 - 28 May
5 - 14 May
4 May
3 May
20 Apr - 2 May
12 -19 Apr
spring
g
12 Jun - 15 Jul
17 Jul - 6 Sep
17 Oct - 28 Dec
8 - 16 Oct
29 Dec 10 - 16 Apr 11
fall
h
15 - 18 May
12 - 14 May
10 - 11 May
9 May
8 May
7 May
1 - 3 Jun
21 - 31 May
20 May
6 May
5 May
4 May
18 Apr - 3 May
17 Apr
spring
50° N
30° N
10° N
160° W
140° W
120° W
100° W
80° W
60° W

A.4 Annual cycle of the remaining inland Swainson's thrushes tracked from pure populations. Dates, stopover sites, breeding and wintering locations are shown for fall (left panels) and spring migration (right panels). Missing dates indicate periods for which the location of the bird could not be estimated. Dashed lines link locations where latitude could not be estimated around the equinox periods. Panels a and b correspond with routes shown in yellow on Figure 2.1, c and d with routes shown in green, e and f with routes shown in orange.


a
15 Jun - 28 Aug
30 Aug - 4 Sep
5 - 6 Sep
8 Oct 10 - 10 Apr 11
fall
b
21 May - 12 Jun
19 - 20 May
18 May
17 May
16 May
15 May
1 - 14 May
18 - 30 Apr
11 - 17 Apr
spring
c
14 Jun - 30 Jul
31 Jul - 6 Sep
8 - 24 Oct
25 Oct
26 Oct
27 Oct
3 Nov 10
- 5 Apr 11
28 Oct - 2 Nov
fall
d
6 -14 Jun
1 - 5 Jun
31 May
27 - 29 May
12 - 26 May
8 - 9 May
5 - 7 May
14 - 29 Apr
spring
e
15 Jun - 31 Jul
1 Aug - 6 Sep
8 - 13 Oct
16 Oct
17 Oct 10 - 7 Apr 11
fall
f
22 May - 13 Jun
21 May
20 May
19 May
18 May
17 May
4 - 16 May
30 Apr - 3 May
28 May
18 - 27 Apr
10 - 17 Apr
spring
50° N
30° N
10° N
160° W
140° W
120° W
100° W
80° W
60° W

Appendix B Supplementary material for Chapter 3

B.1 Additional information on genotyping performed. We screened birds at two introns on the Z chromosome, one in the CHD1Z gene (Fridolfsson and Ellegren 1999), the other in the ADAMT56 gene (Backström et al. 2006). The CHD1Z gene is located between nucleotide numbers 24736734 and 24736133 in the zebra finch genome; ADAMT56 is located between 50815125 and 50815666 (genome.ucsc.edu, assembly July 2008).

We amplified fragments from ten males, five from each taxon sampled in allopatry (coastal: Vancouver; inland: Wisconsin and Illinois) using the primers 2550F (5'-GTTACTGATTCGTCTACGAGA-3') and 2718R (5'-ATTGAAATGATCCAGTGCTTG-3') for CHD1Z and ADAMT56F (5'-GGAGAGAATGGATTTCTGCC-3') and ADAMT56R (5'-TGATTCCAGTCTAGGAAACG-3') for ADAMT56. PCR reactions included 1x PCR buffer (Invitrogen), 1.5 mM MgCl2 (Invitrogen), 0.2 mM dNTP mix (New England Biolabs), 0.5 μM forward and reverse primer, 0.04 units/μl Taq DNA polymerase (New England Biolabs), and 0.025 ng/μl genomic template DNA, in a total volume of 10μl. The thermal cycling profile was 3 minutes at 94ºC followed by 35 cycles of 30 seconds at 94ºC, 30 seconds at 58ºC, and 45 seconds at 72ºC, ending with 5 minutes at 72ºC. Amplified fragments were sequenced bidirectionally by Genome Quebec Innovation Center, using an Applied Biosystem's 3730xl DNA Analyzer.

A 680 bp region of the CHD1Z gene was amplified. There was an A/G polymorphism 220 bp from the 5' end of the fragment. All of the coastal thrushes were homozygous for A in this position; all of the inland thrushes were homozygous for G. The sequence of coastal thrushes contains a cut site for the restriction enzyme *Xab*I while the sequence of inland thrushes does not. A 620bp region of the ADAMT56 gene was amplified. There were two T/C polymorphisms

200 and 480 bp from 5' end of the fragment. All of the inland thrushes were homozygous for T in both positions; all of the coastal thrushes were homozygous for C in both. The sequence of inland thrushes contains a cut site for the restriction enzyme *Psi*I at 480 bp while the sequence of coastal thrushes does not.

Using the same PCR reactions described above, we amplified the introns and digested 2µl of the PCR product with 2 units of the restriction enzyme (*Xab*I and *Psi*I) in its appropriate buffer (New England Biolabs) and in a total volume of 6µl. Products were digested for 2 hours at 37ºC, and digested DNA was visualized by electrophoresis on 2% agarose gel stained with SYBRSafe (Invitrogen). Digestion of CHD1Z cuts the coastal PCR product in two fragments (220 and 480 bp); the inland PCR product remains intact. An individual carrying two copies of the coastal sequence shows two bands; an individual carrying two copies of the inland sequence shows one band; heterozygous individuals produce both cut and intact PCR products generating 3 bands. Digestion of ADAMT56 cuts the inland PCR product into two fragments (130 and 480 bp); the coastal PCR product remains intact. An individual carrying two copies of the inland sequence shows two bands, an individual carrying two copies of the coastal sequence shows one band, heterozygous individuals produce both cut and intact PCR products generating 3 bands. Only males were used in this analysis. Original sequences were obtained from blood samples of birds captured in Vancouver during the breeding season (allopatric coastal site) and tissue samples from the Field Museum of Chicago. These tissue samples were collected in May (one in fall) in Wisconsin and Illinois (allopatric inland sites). We confirmed our PCR-RFLP analysis by assaying birds at allopatric sites, including the Sunshine Coast and Vancouver (allopatric coastal sites, n = 19 for CHD1Z and 17 for ADAMT56) and William's Lake, Kamloops and Kelowna

(allopatric inland sites, $n = 36$ for CHD1Z and 27 for ADAMT56). Each of these birds was homozygous for the corresponding allele.

B.2 Sampling sites used in cline analysis. Location, sample size (_n) and marker frequencies (_freq) are shown. Frequencies were calculated as the number of inland alleles divided by the total number of alleles in each site.

Site	Lat	Long	mtDNA_n	CHD1Z_n	ADAMT56_n	mtDNA_freq	CHD1Z_freq	ADAMT56_freq
Vancouver	49.28	-123.20	21	19	17	0	0	0
Maple Ridge	49.31	-122.48	5	5	5	0	0	0
Chilliwack	49.06	-121.98	16	16	16	0.31	0	0
Cheam	49.22	-121.75	15	15	15	0.13	0.07	0.07
Hope South	49.37	-121.35	19	18	19	0.26	0.19	0.18
Hope North	49.48	-121.25	13	13	13	0.62	0.62	0.62
Coquihalla	49.64	-121.04	13	13	13	1	0.96	0.96
Merrit	49.93	-120.75	18	18	18	0.94	1	1
Kelowna	49.90	-120.13	12	14	14	0.92	1	1

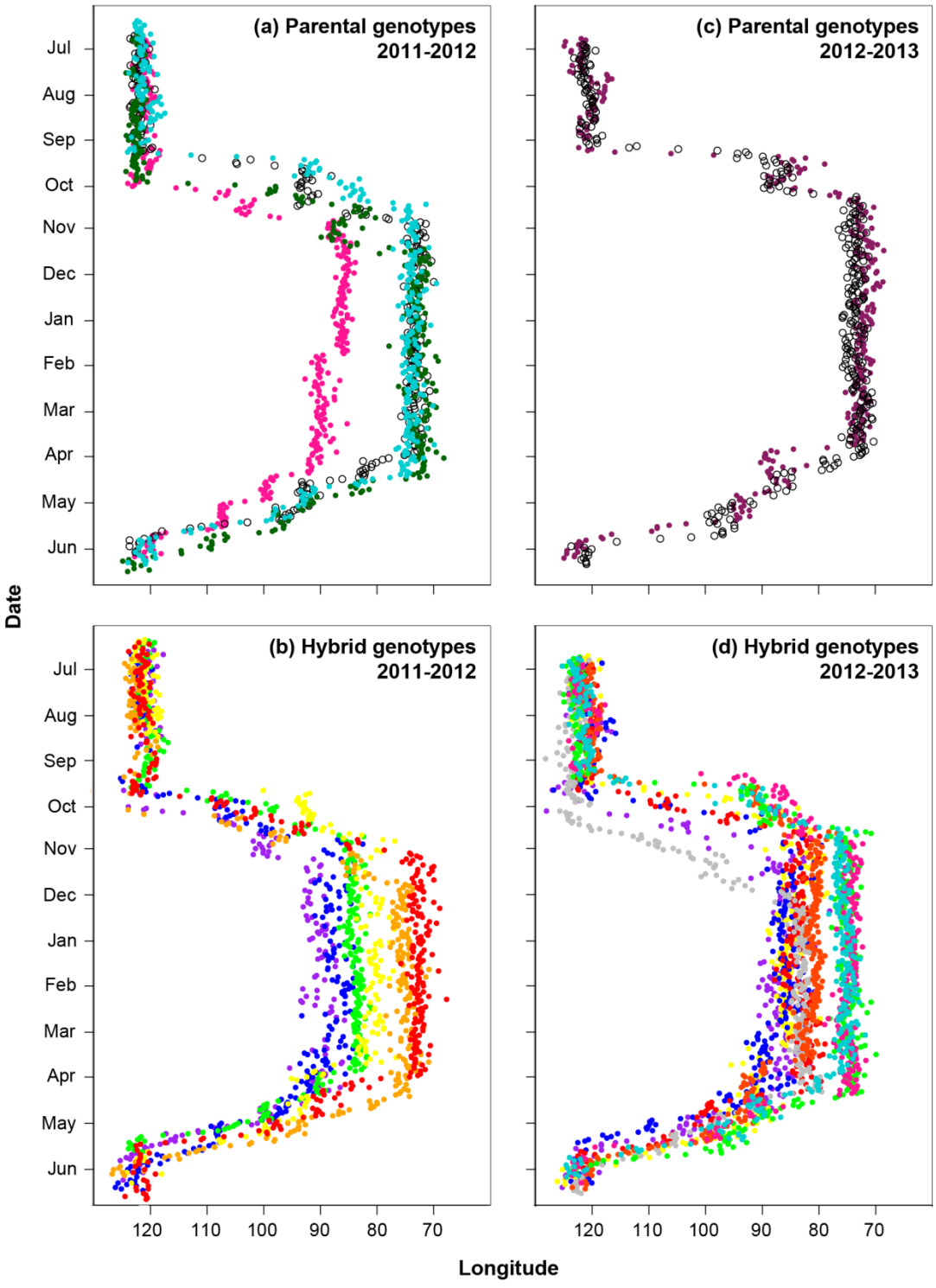
B.3 Details on the fall and spring migration of allopatric (Delmore et al. 2012) and hybrid zone Swainson's thrushes.

Colours correspond to those used in Figures 3.1 and 3.2. Averages for each taxonomic group are shown in bold. Linear regressions between each variable and genetic background are shown; significant results are bolded. Genotypes correspond to alleles from nuclear (CHD1Z and ADAMT56) and mitochondrial markers, respectively. Abbreviations: Geno = genotype (I = inland allele, C = coastal allele, H = inland and coastal alleles), Rec = date of recapture, Loc = location (A = allopatry, H = Hope, P = Pemberton), Dep = date of departure, Arr = date of arrival, Dur = duration of migration in days, Stop = number of days on stopover, Dist = distance in km.

				Fall migration						Spring migration						
	Geno	Rec	Loc	Dep	Arr	Dur	Stop	Dist	Pace	Dep	Arr	Dur	Stop	Dist	Pace	Figure colour
Coastal	C/C/C	2011	A	15-Jul	17-Oct	94	83	3470	37	16-Apr	01-Jun	46	31	4253	92	
	C/C/C	2011	A	25-Jul	18-Oct	85	54	4366	51	06-Apr	28-May	52	39	4441	85	
	C/C/C	2011	A	28-Jul	14-Oct	78	71	3759	48	18-Apr	23-May	36	21	4564	127	
	C/C/C	2011	A	24-Jul	20-Oct	88	72	4466	51	01-Apr						
	C/C/C	2011	A	26-Jul	02-Oct	68	35	4067	60	11-Apr	17-May	36	28	4041	112	
	C/C/C	2012	H	02-Oct	25-Oct	23	16	5573	242	11-Apr	07-Jun	57	43	5559	98	pink
				04-Aug	**16-Oct**	**73**	**55**	**4284**	**82**	**10-Apr**	**27-May**	**45**	**32**	**4572**	**103**	
Inland	I/I/I	2011	A	31-Jul	27-Nov	119	98	5468	46	26-Mar	09-Jun	75	57	5552	74	
	I/I/I	2011	A	30-Jul	03-Nov	96	85	6081	63	04-Apr	06-Jun	63	38	6333	101	
	I/I/I	2011	A	31-Jul	17-Oct	78	66	5519	71	07-Apr	22-May	45	31	6083	135	
	I/I/I	2011	A	28-Aug	08-Oct	41	32	5122	125	10-Apr	21-May	41	32	5532	135	
	I/I/I	2012	P	29-Sep	13-Nov	45	37	7967	177	15-Apr	06-Jun	52	43	7495	144	d. green

				Fall migration						Spring migration						
	Geno	Rec	Loc	Dep	Arr	Dur	Stop	Dist	Pace	Dep	Arr	Dur	Stop	Dist	Pace	Figure colour
	I/I/I	2012	P	11-Sep	25-Oct	44	35	6339	144	30-Mar	25-May	56	40	7708	138	white
	I/I/I	2012	H	05-Aug	13-Oct	69	62	6379	92	14-Apr	24-May	40	29	7390	185	l. blue
	I/I/I	2013	P	18-Jul	10-Oct	84	73	6489	77	24-Mar	27-May	64	53	7221	113	maroon
	I/I/I	2013	H	28-Aug	09-Oct	40	34	7393	185	01-Apr	31-May	60	49	7848	131	white
				16-Aug	**24-Oct**	**68**	**58**	**6306**	**109**	**04-Apr**	**29-May**	**55**	**41**	**6796**	**128**	
Hybrid	H/C/I	2012	P	08-Aug	24-Oct	77	55	5072	66	12-Apr	08-Jun	57	42	6077	107	blue
	H/H/C	2012	H	08-Aug	06-Nov	90	77	4755	53	18-Apr	29-May	41	32	5537	135	purple
	H/H/C	2012	H	08-Aug	22-Oct	75	63	5438	73	27-Mar	20-May	54	42	6697	124	green
	H/H/I	2012	H	11-Jul	24-Oct	105	64	4798	46	22-Mar	20-May	59	44	6390	108	yellow
	H/C/I	2012	P	12-Aug	21-Nov	101	83	7352	73	13-Apr	27-May	44	33	7020	160	orange
	H/H/I	2012	H	02-Aug	02-Nov	92	79	7129	77	31-Mar	23-May	53	44	7256	137	red
	H/H/C	2013	P	24-Jul	25-Oct	93	86	6857	74	13-Apr	25-May	42	26	7254	173	pink
	H/H/C	2013	P	05-Aug	28-Nov	115	104	6766	59	13-Apr	05-Jun	53	39	7321	138	gray
	H/H/I	2013	P	03-Aug	17-Oct	75	58	7671	102	30-Mar	25-May	56	44	7428	133	l. blue
	I/I/C	2013	H	02-Aug	23-Oct	82	77	6470	79	21-Mar	27-May	67	59	6478	97	orange
	H/C/C	2013	H	03-Aug	04-Nov	93	72	5425	58	23-Apr	03-Jun	41	30	5512	134	purple
	C/C/I	2013	H	22-Aug	16-Oct	55	45	5626	102	12-Apr	28-May	46	32	7083	154	red
	C/C/I	2013	P	05-Aug	11-Oct	67	57	5448	81	11-Mar	30-May	80	74	5773	72	yellow
	H/H/I	2013	P	14-Sep	16-Oct	32	26	7558	236	15-Apr	31-May	46	37	7087	154	green
	H/H/C	2013	H	04-Aug	23-Oct	80	72	5742	72	10-Apr	26-May	46	39	6024	131	dark blue
				06-Aug	**27-Oct**	**82**	**68**	**6140**	**83**	**05-Apr**	**28-May**	**52**	**41**	**6596**	**130**	
R^2				0.04	0.01	0.01	0.00	0.40	0.14	0.07	0.00	0.12	0.14	0.45	0.05	
F				1.27	0.39	0.34	0.03	18.38	4.36	2.20	0.01	3.75	4.24	22.11	1.49	
P				0.27	0.53	0.56	0.85	**0.0002**	**0.046**	0.15	0.92	**0.06**	**0.049**	**<0.0001**	0.23	

B.4 Change in longitude over the year for birds tracked from the hybrid zone. Daily estimates of longitude (°W) for birds tracked between 2011 and 2012 (ab) and 2012 and 2013 (cd). Panels a and c show birds genotyped as parental from the hybrid zone; panels b and d birds genotyped as hybrid. Colours correspond to those used in Figures 3.1 and 3.2.


(a) Parental genotypes
2011-2012
(c) Parental genotypes
2012-2013
(b) Hybrid genotypes
2011-2012
(d) Hybrid genotypes
2012-2013
Date
Jul
Aug
Sep
Oct
Nov
Dec
Jan
Feb
Mar
Apr
May
Jun
120
110
100
90
80
70
Longitude

B.5 Daily location estimates for birds tracked from the hybrid zone. Daily estimates of longitude and latitude for Swainson's thrushes tracked in this study, including standard deviations (SD), sample sizes (*n*) and standard errors (SE) for days when birds were stationary (i.e. on the breeding or wintering grounds, at stopover sites). Colours correspond to those used in Figures 3.1 and 3.2. Data from birds tracked between 2011 and 2012 are shown first (Figure 3.1ab and Figure 3.2ab) and followed by data from birds tracked between 2012 and 2013 (Figure 3.1cd and Figure 3.2cd). Missing values indicate periods for which the location of the bird could not be estimated accurately (e.g., around the equinoxes).

Date	Longitude (°W)	SD	*n*	SE	Latitude (°N)	SD	*n*	SE
Birds tracked between 2011 and 2012								
White								
Jun 22 - Sep 11	122.05	0.92	63	0.12	49.94	1.29	56	0.17
Sep 13	110.89		1					
Sep 14 - 18	102.52	3.18	4	1.59				
Sep 20 - Oct 15	92.44	1.38	21	0.30	36.18	5.89	5	2.63
Oct 17 - 22	83.60	1.53	6	0.62	22.75	2.74	6	1.12
Oct 23 - 24	78.18	0.33	2	0.23	20.84	4.72	2	3.34
Oct 25 11 - Mar 30 12	73.21	1.05	134	0.09	11.18	8.14	115	0.76
Apr 1 - 2	78.49	0.76	2	0.53				
Apr 3 - 14	82.02	0.95	11	0.29	8.29	12.42	10	3.93
Apr 15	86.15		1		11.43		1	
Apr 16 - May 3	92.25	1.34	18	0.32	14.65	5.66	18	1.33
May 4 - 13	96.23	1.30	10	0.41	33.38	3.28	10	1.04
May 14	102.91		1		41.21		1	
May 15	106.91		1		44.30		1	
May 17	111.02		1		45.46		1	
May 18	113.01		1		44.04		1	
May 20	117.98		1		45.60		1	
May 23 - 24	119.28	0.16	2	0.11	49.85	0.17	2	0.12
May 25 - Jun 3	122.50	0.95	7	0.36	49.48	1.01	7	0.38

Date	Longitude (°W)	SD	*n*	SE	Latitude (°N)	SD	*n*	SE
Green								
Jun 13 - Aug 8	121.11	0.88	48	0.13	49.98	0.94	48	0.14
Aug 12 - Sep 20	120.25	1.59	32	0.28	46.10	1.67	21	0.36
Sep 21	113.60		1					
Sep 22 - 28	107.61	1.43	7	0.54				
Sep 29	102.79		1					
Sep 30 - Oct 6	98.95	1.02	6	0.42				
Oct 8 - 20	92.49	1.27	8	0.45	29.66	9.01	8	3.19
Oct 21	86.71		1		15.62		1	
Oct 22 11 - Mar 27 12	83.80	1.09	130	0.10	15.44	3.92	111	0.37
Mar 28 - Apr 11	89.85	0.97	13	0.27	13.69	3.59	7	1.36
Apr 12 - 15	93.13	0.35	4	0.18	15.56	3.24	4	1.62
Apr 16 - 29	99.56	0.86	12	0.25	16.80	2.15	12	0.62
Apr 30 - May 4	104.79	1.26	5	0.56	21.17	1.72	5	0.77
May 5	106.96		1		28.22		1	
May 6	108.10		1		30.82		1	
May 7	110.74		1		32.70		1	
May 8	113.38		1		34.89		1	
May 9	117.77		1		36.88		1	
May 10 - 12	121.95	0.32	3	0.19	37.19	1.72	3	0.99
May 13	121.04		1		42.37		1	
May 14 - 19	123.31	0.48	6	0.19	45.13	1.62	6	0.66
May 20	121.48		1		50.55		1	
Yellow								
Jun 12 - Jul 11	121.39	1.34	25	0.27	49.14	0.87	25	0.17
Jul 12 - Aug 16	120.00	1.13	31	0.20	47.28	0.93	31	0.17
Sep 21	100.10		1					
Sep 23 - Oct 19	92.00	1.80	21	0.39	39.01	15.50	7	5.86
Oct 20 - 23	84.51	1.64	3	0.95	32.97	3.14	3	1.81
Oct 24 11 - Mar 22 12	80.49	1.71	81	0.19	22.95	7.69	68	0.93
Mar 23	87.65		1					
Mar 27 - Apr 28	92.74	1.68	18	0.40	15.15	9.38	11	2.83
Apr 29 - May 8	98.93	1.79	7	0.68	23.53	2.15	7	0.81
May 9	99.14		1		29.68		1	
May 10	103.15		1		36.30		1	
May 11 - 14	107.50	0.35	4	0.17	36.65	0.63	4	0.32
May 17	108.77		1		41.99		1	

Date	Longitude (°W)	SD	*n*	SE	Latitude (°N)	SD	*n*	SE
May 20 - Jun 10	121.52	2.03	7	0.77	47.96	1.85	7	0.70
Blue								
Jun 22 - Aug 8	122.05	0.89	36	0.15	50.44	1.08	35	0.18
Aug 23 - Sep 22	122.09	1.80	10	0.57	46.52	0.52	2	0.37
Sep 23 - 25	117.70	1.03	3	0.59				
Sep 26	111.66		1					
Sep 27	108.37		1					
Sep 28 - Oct 11	106.14	1.55	13	0.43	47.47	4.37	2	3.09
Oct 13	103.06		1		40.82		1	
Oct 14 - 21	99.56	0.78	7	0.30	40.11	3.28	7	1.24
Oct 22	95.75		1		35.38		1	
Oct 24 11 - Apr 12 12	88.06	2.65	109	0.25	21.50	13.30	83	1.46
Apr 13 - 28	96.63	1.09	16	0.27	15.94	5.61	16	1.40
Apr 29	100.31		1		17.82		1	
Apr 30	104.21		1		20.12		1	
May 1 - 15	108.12	0.51	14	0.14	26.50	2.12	14	0.57
May 16 - 21	111.94	1.64	6	0.67	31.48	2.34	6	0.95
May 22 - 28	118.69	1.50	7	0.57	33.66	1.68	7	0.64
May 29 - Jun 3	122.40	1.25	6	0.51	39.18	1.68	6	0.68
Jun 6	122.42		1		45.48		1	
Jun 8 - 15	123.92	1.70	8	0.60	49.87	0.98	8	0.35
Purple								
Jun 19 - Aug 8	120.45	1.01	41	0.16	49.51	1.00	41	0.16
Aug 10 - Oct 5	120.40	2.02	16	0.51	46.31	2.21	7	0.84
Oct 6	114.21		1					
Oct 8 - 14	106.56	1.01	6	0.41	45.34	6.38	6	2.60
Oct 20 - 27	99.97	1.51	6	0.62	41.13	4.50	6	1.84
Oct 28 - Nov 5	99.86	1.32	9	0.44	30.38	2.32	9	0.77
Nov 6 11 - Apr 18 12	90.23	1.93	75	0.22	21.03	14.54	61	1.86
Apr 20 - May 2	99.94	1.20	11	0.36	16.95	7.65	11	2.31
May 3 - 5	107.77	1.20	3	0.69	24.60	6.59	3	3.80
May 6 - 7	111.55	0.08	2	0.05	32.96	0.49	2	0.34
May 8 - 15	116.28	1.22	8	0.43	34.97	2.12	8	0.75
May 16 - 26	122.37	1.75	11	0.53	40.55	2.02	11	0.61
May 29 - Jun 8	122.96	1.65	4	0.83	48.94	1.75	4	0.88

Date	Longitude (°W)	SD	*n*	SE	Latitude (°N)	SD	*n*	SE
Pink								
Jun 20 - Oct 2	121.00	1.19	82	0.13	49.79	0.92	55	0.12
Oct 3	115.49		1					
Oct 4	112.06		1					
Oct 5 - 10	107.15	0.97	5	0.44	28.75	11.88	2	8.40
Oct 11 - 21	103.22	1.00	10	0.32	29.72	7.75	8	2.74
Oct 22 - 23	98.08	1.12	2	0.80	19.70	1.75	2	1.24
Oct 25 11 - Jan 22 12	86.20	1.02	89	0.11	15.72	3.78	89	0.40
Jan 23 - Apr 11	90.09	1.36	77	0.15	16.96	6.20	47	0.91
Apr 12	95.09		1		18.44		1	
Apr 13 - 29	99.32	0.71	17	0.17	16.71	2.81	17	0.68
Apr 30	103.34		1		21.80		1	
May 1 - May 18	107.30	0.69	18	0.16	28.45	0.82	18	0.19
May 19	109.87		1		31.25		1	
May 20	113.85		1		34.32		1	
May 21 - 27	119.01	1.09	7	0.41	36.10	0.87	7	0.33
May 28	119.65		1		40.03		1	
May 29	120.62		1		42.69		1	
May 30	120.58		1		44.81		1	
May 31 - Jun 3	120.27	0.65	4	0.32	46.54	1.79	4	0.90
Jun 4 - 6	120.75	2.38	3	1.38	48.35	0.29	3	0.17
Jun 7 - 9	120.95	0.11	3	0.06	49.95	0.69	2	0.49
Dark Green								
Jun 21 - Sep 29	122.34	0.88	77	0.10	50.00	1.11	52	0.15
Sep 30	113.37		1					
Oct 1	105.08		1					
Oct 2 - 7	98.45	0.60	6	0.25				
Oct 8 - 13	93.27	1.05	6	0.43	47.74	6.17	6	2.52
Oct 14 - 23	83.10	2.09	9	0.70	45.03	6.45	9	2.15
Oct 24 - Nov 11	86.42	1.90	18	0.45	30.64	4.47	18	1.05
Nov 12	77.84		1		8.03		1	
Nov 13 11 - Apr 15 12	72.53	1.58	127	0.14	11.70	12.04	98	1.22
Apr 16 - Apr 25	83.13	1.29	10	0.41	8.47	6.43	10	2.03
Apr 26	87.32		1		12.74		1	
Apr 27 - May 6	92.15	0.92	9	0.31	14.68	4.31	9	1.44
May 7 - 20	97.08	0.96	11	0.29	30.23	2.30	11	0.69

Date	Longitude (°W)	SD	*n*	SE	Latitude (°N)	SD	*n*	SE
May 21	100.84		1		42.04		1	
May 22 - Jun 1	108.79	1.48	11	0.45	44.58	1.70	11	0.51
Jun 2 - 4	114.55	0.06	2	0.04	46.76	1.12	2	0.79
Jun 6 - 16	122.23	2.00	10	0.63	50.31	1.24	10	0.39
Orange								
Jun 25 - Aug 12	122.94	0.72	41	0.11	51.14	0.66	41	0.10
Aug 19 - Oct 7	122.05	3.09	15	0.80	45.79	2.27	9	0.76
Oct 8 - 17	105.76	1.78	8	0.63	41.38	11.34	8	4.01
Oct 18 - 27	97.49	1.06	9	0.35	27.24	5.29	9	1.76
Nov 2 - 18	84.03	1.29	13	0.36	22.68	6.83	13	1.89
Nov 19	80.53		1		9.15		1	
Nov 21 11 - Ap 13 12	75.08	1.49	131	0.13	10.41	9.08	102	0.90
Apr 15 - 21	80.39	0.89	7	0.34	7.55	6.99	7	2.64
Apr 22 - 26	86.44	1.08	5	0.48	11.73	2.87	5	1.29
Apr 27 - May 9	93.90	1.29	12	0.37	15.91	2.38	12	0.69
May 11 - 13	98.58	0.83	3	0.48	23.37	0.16	3	0.09
May 14	102.16		1		25.38		1	
May 15	102.53		1		29.42		1	
May 16	105.90		1		32.56		1	
May 17	108.64		1		35.19		1	
May 18 - 20	111.37	0.68	3	0.39	38.58	0.67	3	0.39
May 21 - 25	115.49	1.67	5	0.75	43.20	1.35	5	0.60
May 27 - Jun 5	124.44	1.92	9	0.64	47.53	1.06	9	0.35
Light Blue								
Jun 13 - Aug 5	121.43	1.01	52	0.14	50.47	0.93	52	0.13
Aug 6 - Sep 9	120.62	1.73	28	0.33	47.97	1.92	23	0.40
Sep 10 - 13	108.67	8.89	4	4.45				
Sep 14 - 25	91.87	1.33	12	0.39				
Sep 26	88.41		1					
Sep 27 - Oct 12	84.13	1.51	15	0.39	15.67	16.76	5	7.49
Oct 13 11 - Apr 14 12	74.00	1.31	178	0.10	11.34	7.04	148	0.58
Apr 15 - 19	82.56	1.19	5	0.53	8.34	7.58	5	3.39
Apr 20 - May 3	92.25	1.43	14	0.38	14.62	5.97	14	1.59
May 4 - 12	97.28	1.20	7	0.46	28.89	2.10	7	0.79
May 13	99.54		1		37.13		1	
May 14	104.16		1		36.57		1	

Date	Longitude (°W)	SD	*n*	SE	Latitude (°N)	SD	*n*	SE
May 16 - 17	108.90	0.36	2	0.26	43.81	1.23	2	0.87
May 18 - 21	112.77	1.23	4	0.61	48.41	0.75	4	0.38
May 24 - Jun 11	120.92	1.31	19	0.30	49.75	0.84	19	0.19
Red								
Jun 14 - Aug 2	121.59	0.88	43	0.13	49.76	1.01	43	0.15
Aug 4 - Sep 24	121.28	1.58	31	0.28	45.29	1.75	11	0.53
Sep 25	114.83		1					
Sep 26 - Oct 1	107.98	1.11	6	0.45				
Oct 2- 12	101.74	3.18	11	0.96	40.45	7.44	5	3.33
Oct 13 - 20	93.56	0.87	8	0.31	36.24	7.53	8	2.66
Oct 24 - 30	84.48	0.90	6	0.37	15.86	2.87	6	1.17
Oct 31	78.71		1		9.42		1	
Nov 2 11 - Mar 31 12	72.62	1.27	144	0.11	11.74	4.86	117	0.45
Apr 1 - 8	83.20	2.21	7	0.84	8.78	13.86	5	6.20
Apr 9 - 27	89.84	1.73	16	0.43	13.71	10.23	16	2.56
Apr 29 - 30	98.39	0.95	2	0.68	23.96	1.44	2	1.02
May 1 - 8	98.51	1.32	8	0.47	31.63	1.71	8	0.60
May 9 - 10	105.53	2.13	2	1.51	37.42	0.01	2	0.01
May 11 - 13	111.91	1.85	3	1.07	41.50	0.84	3	0.49
May 14	117.41		1		43.92		1	
May 15 - 22	121.74	0.92	5	0.41	45.96	0.47	5	0.21
May 23 - Jun 25	121.47	1.11	24	0.23	48.54	0.96	24	0.19
Birds tracked between 2012 and 2013								
White								
Jun 29 - Aug 28	120.94	0.96	54	0.13	50.43	1.06	53	0.15
Aug 29 - Sep 04	119.70	1.04	7	0.39	45.40	3.22	7	1.22
Sep 05 - 06	112.80	0.95	2	0.67	42.37		1	
Sep 07	104.80		1					
Sep 08	97.89		1					
Sep 09 - Oct 06	88.23	2.74	28	0.52				
Oct 07	80.96		1		42.87		1	
Oct 09 12 - Apr 1 13	73.40	1.27	166	0.10	11.27	5.11	138	0.43
Apr 2 - 10	78.17	1.14	8	0.40	7.99	14.12	5	6.31
Apr 11 - May 1	87.07	2.26	11	0.68	12.46	9.56	11	2.88
May 2 - 14	96.62	2.15	12	0.62	15.79	8.57	12	2.47

Date	Longitude (°W)	SD	*n*	SE	Latitude (°N)	SD	*n*	SE
May 15 - 23	96.84	1.57	7	0.59	28.50	5.78	8	2.04
May 24	98.40		1		42.26		1	
May 25	102.50		1		47.25		1	
May 26	107.97		1		49.15		1	
May 28	115.66		1		50.44		1	
May 31 - Jun 13	121.30	0.92	11	0.28	50.03	0.89	11	0.27
Green								
Jun 24 - Sep 14	122.64	0.98	72	0.12	50.99	1.36	64	0.17
Sep 15	115.89		1					
Sep 16	108.97		1					
Sep 17	102.06		1					
Sep 18 - Oct 12	90.55	2.36	23	0.49	35.32	5.89	5	2.63
Oct-13	90.47		1		18.16		1	
Oct-14	88.53		1		13.31		1	
Oct-15	85.21		1		12.36		1	
Oct 16 12 - Apr 15 13	74.76	1.37	173	0.10	8.41	10.40	143	0.87
Apr 16 - Apr 21	83.40	1.22	6	0.50	8.81	4.52	6	1.85
Apr 22 - May 1	89.46	1.40	10	0.44	13.57	3.01	10	0.95
May 2 - 18	95.45	1.44	16	0.36	29.80	3.31	16	0.83
May 19 - 22	98.41	1.36	4	0.68	38.38	0.98	4	0.49
May 23	103.55		1		42.08		1	
May 24	106.78		1		44.84		1	
May 25 - 26	112.55	0.24	2	0.17	47.62	0.69	2	0.49
May 27 - 30	118.33	1.91	4	0.95	49.53	0.72	4	0.36
May 31 - Jun 6	123.02	0.33	6	0.14	51.01	1.42	6	0.58
Blue								
Jun 27 - Aug 4	121.74	1.09	31	0.20	49.48	1.20	31	0.22
Aug 5 - Sep 28	120.05	2.73	32	0.48	43.74	2.17	14	0.58
Sep 29 - Oct 5	106.94	0.75	4	0.37				
Oct 6	100.27		1					
Oct 7 - 19	91.70	1.16	5	0.52	39.38	5.14	2	3.64
Oct 23 12 - Feb 27 13	84.90	2.09	109	0.20	15.90	8.64	109	0.83
Mar 03 - Apr 10	89.36	1.85	37	0.30	13.60	24.41	6	9.96
Apr 11 - 26	95.26	1.30	14	0.35	16.30	6.51	14	1.74
Apr 27 - 28	100.88	1.97	2	1.40	19.97	0.29	2	0.21
Apr 29	104.42		1		22.74		1	

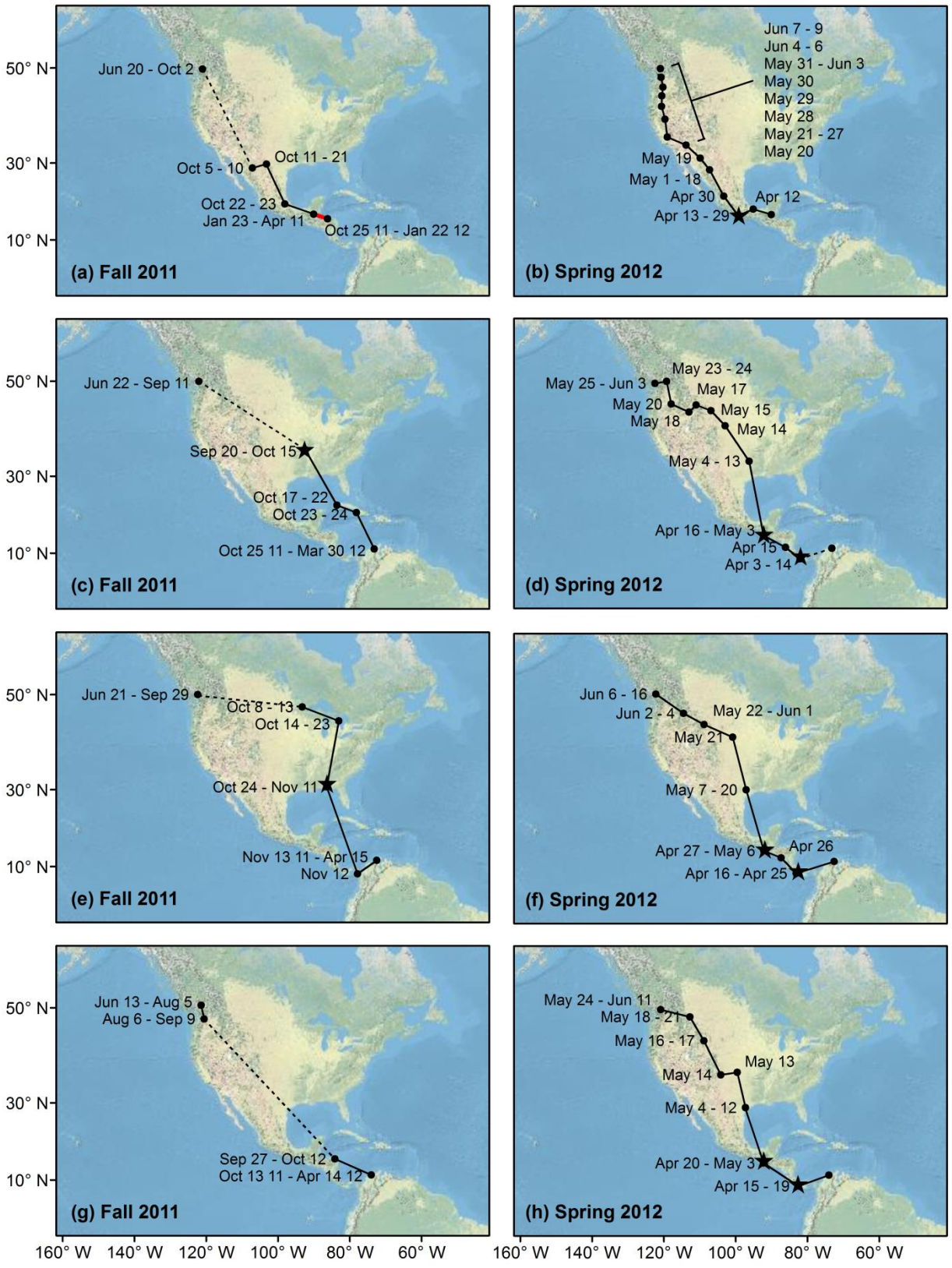
Date	Longitude (°W)	SD	*n*	SE	Latitude (°N)	SD	*n*	SE
Apr 30 - May 8	110.00	0.87	9	0.29	28.31	2.04	9	0.68
May 9 - 12	116.94	1.12	4	0.56	31.87	1.52	4	0.76
May 13 - 14	120.98	0.45	2	0.32	35.01	1.39	2	0.98
May 15 - 25	121.92	0.73	3	0.42	44.17	1.75	3	1.01
May 26 - Jun 10	122.08	1.54	14	0.41	48.05	1.37	14	0.37
Pink								
June 25 - July 24	122.69	0.66	38	0.11	51.48	0.72	28	0.14
July 25 - Sep 7	120.85	1.97	20	0.44	45.80	2.18	17	0.53
Sep 10	100.81		1					
Sep 11 - Oct 22	86.43	3.76	38	0.61	36.16	3.55	13	0.98
Oct 23	81.43		1		19.73		1	
Oct 25 12 - Apr 13 13	74.63	1.20	156	0.10	10.42	11.74	127	1.04
Apr 15 - 28	90.64	1.85	13	0.51	14.58	9.31	13	2.58
Apr 29 - May 12	97.62	0.98	11	0.30	30.86	0.36	11	0.11
May 13	99.79		1		39.45		1	
May 14	102.41		1		42.68		1	
May 15	107.53		1		46.56		1	
May 17	116.14		1		47.31		1	
May 18	119.51		1		45.31		1	
May 25 - Jun 11	122.75	1.10	12	0.32	49.46	1.03	12	0.30
Gray								
Jun 26 - Aug 5	123.07	1.07	34	0.18	50.69	1.11	34	0.19
Aug 7 - Oct 19	123.16	1.95	61	0.25	44.99	5.54	36	0.92
Oct 20 - Oct 24	112.43	1.47	5	0.66	29.37	4.19	5	1.87
Oct 25 - Oct 29	108.89	0.69	5	0.31	27.79	3.42	5	1.53
Oct 30 - Nov 02	105.22	1.88	4	0.94	23.71	2.16	4	1.08
Nov 3 - 26	98.00	2.62	18	0.62	28.37	3.73	18	0.88
Nov 27	91.78		1		18.45		1	
Nov 28 12 - Apr 13 13	83.12	1.36	129	0.12	14.43	11.33	101	1.13
Apr 14 - 24	91.03	1.17	11	0.35	14.36	4.95	11	1.49
Apr 25	94.65		1		17.67		1	
Apr 26 - May 07	98.23	0.77	12	0.22	16.61	4.43	12	1.28
May 8	102.75		1		19.22		1	
May 9 - 17	105.57	0.39	9	0.13	20.97	3.08	9	1.03
May 18 - 19	108.63	0.17	2	0.12	28.00	0.49	2	0.35

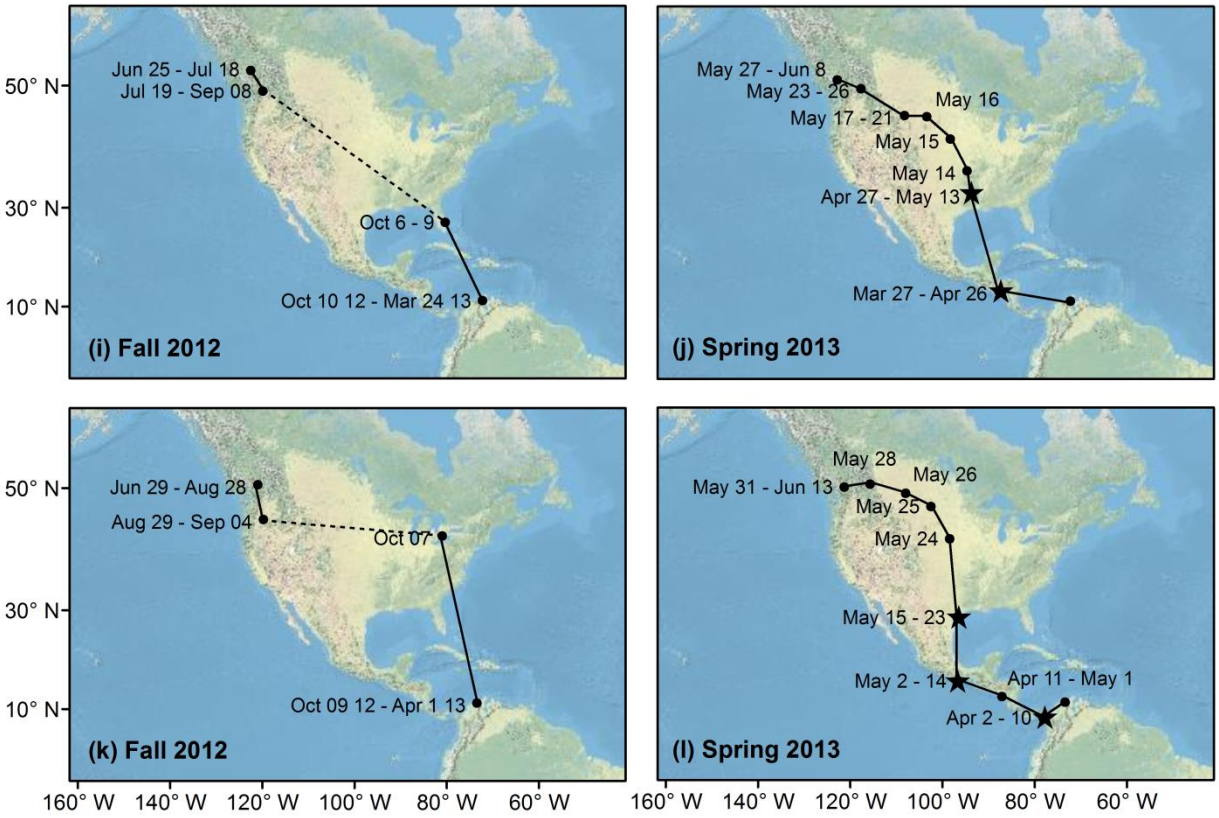
Date	Longitude (°W)	SD	*n*	SE	Latitude (°N)	SD	*n*	SE
May 20 - 22	108.34	0.20	2	0.14	35.63	0.42	2	0.30
May 23 - 25	112.11	0.62	3	0.36	34.71	1.33	3	0.77
May 26	114.60		1		39.05		1	
May 28 - 30	118.42	0.17	3	0.10	36.42	1.41	3	0.81
May 31 - Jun 3	122.46	1.65	3	0.95	40.43	2.34	3	1.35
Jun 5 - 18	122.62	0.72	13	0.20	49.74	1.43	13	0.40
Light Blue								
Jun 25 - Aug 3	122.17	1.20	36	0.20	50.04	1.12	36	0.19
Aug 8 - Sep 11	121.33	0.99	31	0.18	47.12	1.97	26	0.39
Sep 12 - 17	115.14	1.31	6	0.54				
Sep 18	98.40		1					
Sep 19	93.49		1					
Sep 20 - Oct 06	90.84	2.73	15	0.70				
Oct 07 - 08	86.94	1.10	2	0.77	40.52	12.37	2	8.74
Oct 12 - 14	82.89	0.54	3	0.31	33.50	4.08	3	2.36
Oct 15	84.59		1		10.02		1	
Oct 16	80.01		1		9.01		1	
Oct 17 12 - Mar 30	75.57	1.17	160	0.09	7.42	5.65	137	0.48
Mar 31 - Apr 11 13	80.13	1.00	12	0.29	7.71	6.67	7	2.52
Apr 12 -13	85.55	0.66	2	0.47	12.11	0.16	2	0.11
Apr 14 - 30	90.60	1.43	16	0.36	14.08	8.65	16	2.16
May 01 - 07	98.16	1.63	7	0.62	26.38	2.38	7	0.90
May 08 - 13	94.67	1.46	6	0.60	31.32	2.14	6	0.88
May 16	102.65		1		42.63		1	
May 17	107.27		1		43.66		1	
May 18 - 23	112.80	1.46	5	0.65	43.70	2.24	5	1.00
May 25 - June 07	123.06	1.21	13	0.34	50.24	1.09	13	0.30
Orange								
Jun 29 - Aug 2	120.78	1.15	28	0.22	50.18	1.00	28	0.19
Aug 3 - Sep 11	119.71	0.87	29	0.16	47.13	2.52	23	0.52
Sep 12 - 16	114.54	1.94	4	0.97				
Sep 17	100.19		1					
Sep 18 - Oct 1	91.28	0.83	14	0.22				
Oct 2 - 15	87.43	1.36	13	0.38	35.21	8.00	7	3.02
Oct 16 - 22	81.77	3.39	7	1.28	29.90	7.77	5	3.47
Oct 23 12 - Mar 21 13	80.99	1.16	148	0.10	8.81	9.08	132	0.79

Date	Longitude (°W)	SD	*n*	SE	Latitude (°N)	SD	*n*	SE
Mar 22 - May 3	92.04	1.79	41	0.28	14.66	6.17	28	1.17
May 4 - 5	97.20	1.43	2	1.01	18.33	1.30	2	0.92
May 6 - 7	100.24	0.70	2	0.49	25.56	1.89	2	1.34
May 8 - 15	99.97	0.92	8	0.33	35.32	3.19	8	1.13
May 16 - 19	108.70	1.38	4	0.69	39.80	0.86	4	0.43
May 20 - 23	113.92	1.17	4	0.59	40.51	0.75	4	0.38
May 24 - 26	118.42	1.10	3	0.64	44.23	1.17	3	0.67
May 27 - Jun 15	121.04	1.02	14	0.27	48.90	1.16	14	0.31
Purple								
Jun 28 - Aug 3	121.64	1.02	24	0.21	49.42	1.28	24	0.26
Aug 7 - Oct 7	121.49	3.20	23	0.67	47.52	2.06	11	0.62
Oct 8	-113.91		1					
Oct 11 - 19	104.76	1.90	7	0.72	47.14	4.12	3	2.38
Oct 21 - 30	96.55	3.77	4	1.89	42.91	6.74	4	3.37
Nov 4 12 - Apr 23 13	86.72	2.44	132	0.21	15.79	12.44	106	1.21
Apr 24 - May 11	102.64	2.22	17	0.54	19.11	3.77	17	0.91
May 13 - 26	118.53	4.23	11	1.27	36.78	1.99	11	0.60
Jun 3 - 14	122.29	2.72	5	1.22	47.98	1.81	5	0.81
Red								
Jun 29 - Aug 22	120.89	1.12	51	0.16	50.78	0.92	51	0.13
Aug 25 - Sep 21	121.29	1.28	19	0.29	47.95	2.11	6	0.86
Sep 22	115.38		1					
Sep 23	113.72		1					
Sep 24 - Oct 5	107.09	1.93	11	0.58				
Oct 7 - 13	96.18	1.42	8	0.50	28.75	5.29	7	2.00
Oct 14 -15	89.88	0.40	2	0.28	34.37	1.46	2	1.04
Oct 16 12 - Apr 12 13	83.23	1.83	153	0.15	23.05	12.10	126	1.08
Apr 13 - 16	90.92	0.69	4	0.34	14.05	3.80	4	1.90
Apr 17	95.75		1		18.81		1	
Apr 18 - May 8	99.56	1.09	20	0.24	17.00	3.97	20	0.89
May 9 - 16	106.66	1.04	8	0.37	26.73	1.31	8	0.46
May 17	110.02		1		31.25		1	
May 18	113.01		1		33.59		1	
May 19	114.62		1		37.08		1	
May 21 - 23	120.94	1.25	3	0.72	36.99	1.79	3	1.03
May 24	122.53		1		42.93		1	

Date	Longitude (°W)	SD	*n*	SE	Latitude (°N)	SD	*n*	SE
May 28 - Jun 02	121.20	1.12	6	0.46	48.87	0.97	6	0.40
Yellow								
Jun 23 - Aug 5	121.56	1.41	39	0.23	50.33	1.45	39	0.23
Aug 7 - Sep 17	121.22	1.39	24	0.28	43.26	2.67	15	0.69
Sep 19 - 21	110.87	0.05	3	0.03				
Sep 22	107.88		1					
Sep 23	103.22		1					
Sep 24 - Oct 08	91.69	4.47	14	1.20	18.70	6.59	2	4.66
Oct 11 12 - Mar 11 13	84.97	2.21	136	0.19	15.91	9.91	130	0.87
Mar 12 - May 10	92.17	2.57	38	0.42	14.66	14.73	28	2.78
May 11	98.28		1		23.18		1	
May 12 - 15	102.29	1.81	4	0.90	35.71	2.69	4	1.34
May 16 - 24	109.36	2.11	9	0.70	40.06	1.63	9	0.54
May 25 - 29	118.07	1.54	5	0.69	42.76	1.10	5	0.49
May 30 - June 11	123.68	1.53	12	0.44	49.09	1.26	12	0.36
Maroon								
Jun 25 - Jul 18	122.44	1.10	24	0.22	52.05	0.68	24	0.14
Jul 19 - Sep 08	119.84	1.59	40	0.25	49.24	2.10	37	0.34
Sep 09	115.98		1					
Sep 10	106.06		1					
Sep 11	98.53		1					
Sep 12 - Oct 4	85.37	2.46	23	0.51				
Oct 5	81.44		1					
Oct 6 - 9	80.29	0.94	4	0.47	27.23	16.02	3	9.25
Oct 10 12 - Mar 24 13	72.21	1.25	161	0.10	11.26	6.81	142	0.57
Mar 26	78.48		1					
Mar 27 - Apr 26	87.80	2.07	27	0.40	13.45	15.84	18	3.73
Apr 27 - May 13	93.80	0.89	14	0.24	32.60	1.61	14	0.43
May 14	94.66		1		36.76		1	
May 15	98.28		1		42.09		1	
May 16	103.40		1		45.57		1	
May 17 - 21	108.22	1.30	4	0.65	45.74	2.54	4	1.27
May 23 - 26	117.68	0.99	3	0.57	49.66	1.81	3	1.05
May 27 - Jun 8	122.80	1.22	13	0.34	50.87	1.04	13	0.29

B.6 Annual cycle of Swainson's thrushes genotyped as parental from the hybrid zone and showing similar behaviours to allopatric birds (Delmore et al. 2012). Dates, stopover sites, breeding and wintering locations are shown for fall (left panels) and spring migration (right panels). Missing dates indicate periods for which the location of the bird could not be estimated accurately. Dashed lines indicate periods when latitude could not be estimated around the equinox periods. Panels a – b show data from the individual recaptured in 2012, genotyped as coastal and demonstrating the use of two sites on the wintering grounds (linked by the thick red line; i.e., intratropical migration). Panels c – l show data from individuals genotyped as inland and demonstrating the use of different routes to navigate the Gulf of Mexico on fall and spring migration (i.e., loop migration). The location of long-term stopover sites similar to those described in Delmore et al. (2012) are shown with black stars. Panels c – h show data from birds recaptured in 2012; panels i – l show data from birds recaptured in 2013. Note, each row shows fall and spring migration for a single individual (referring to colors used in Figure 2.1, a and b = pink, c and d = white, e and f = dark green, g and h = light blue, i and j = maroon, k and l = white).


50° N
30° N
10° N
Jun 20 - Oct 2
Oct 5 - 10
Oct 11 - 21
Oct 22 - 23
Jan 23 - Apr 11
Oct 25 11 - Jan 22 12
(a) Fall 2011
Jun 7 - 9
Jun 4 - 6
May 31 - Jun 3
May 30
May 29
May 28
May 21 - 27
May 20
May 19
May 1 - 18
Apr 30
Apr 12
Apr 13 - 29
(b) Spring 2012
Jun 22 - Sep 11
Sep 20 - Oct 15
Oct 17 - 22
Oct 23 - 24
Oct 25 11 - Mar 30 12
(c) Fall 2011
May 23 - 24
May 25 - Jun 3
May 17
May 20
May 15
May 18
May 14
May 4 - 13
Apr 16 - May 3
Apr 15
Apr 3 - 14
(d) Spring 2012
Jun 21 - Sep 29
Oct 8 - 13
Oct 14 - 23
Oct 24 - Nov 11
Nov 13 11 - Apr 15
Nov 12
(e) Fall 2011
Jun 6 - 16
Jun 2 - 4
May 22 - Jun 1
May 21
May 7 - 20
Apr 27 - May 6
Apr 26
Apr 16 - Apr 25
(f) Spring 2012
Jun 13 - Aug 5
Aug 6 - Sep 9
Sep 27 - Oct 12
Oct 13 11 - Apr 14 12
(g) Fall 2011
May 24 - Jun 11
May 18 - 21
May 16 - 17
May 13
May 14
May 4 - 12
Apr 20 - May 3
Apr 15 - 19
(h) Spring 2012
160° W
140° W
120° W
100° W
80° W
60° W


Jun 25 - Jul 18
Jul 19 - Sep 08
Oct 6 - 9
Oct 10 12 - Mar 24 13
(i) Fall 2012
May 27 - Jun 8
May 23 - 26
May 17 - 21
May 16
May 15
May 14
Apr 27 - May 13
Mar 27 - Apr 26
(j) Spring 2013
Jun 29 - Aug 28
Aug 29 - Sep 04
Oct 07
Oct 09 12 - Apr 1 13
(k) Fall 2012
May 28
May 31 - Jun 13
May 26
May 25
May 24
May 15 - 23
May 2 - 14
Apr 11 - May 1
Apr 2 - 10
(l) Spring 2013
50° N
30° N
10° N
160° W
140° W
120° W
100° W
80° W
60° W

B.7 Additional analyses examining the relationship between the migratory behaviour of birds tracked and their genetic background. We showed a strong positive relationship between genetic background and migratory routes employed by birds on spring and fall migration at 30°N with and without allopatric birds. A similar relationship was recovered when the analyses were run without parental genotypes (i.e., removing individuals from both allopatric and hybrid sites with parental genotypes; fall migration: $R^2 = 0.36$, $F_{1,13} = 7.36$, $P = 0.02$; spring migration: $R^2 = 0.35$, $F_{1,13} = 7.06$, $P = 0.02$; wintering grounds: $R^2 = 0.25$, $F_{1,13} = 4.30$, $P = 0.06$) and when all of these analyses were run at 40°N (Fall migration: Full dataset $R^2 = 0.47$, $F_{1,28} = 25.12$, $P < 0.0001$, allopatric sites removed $R^2 = 0.24$, $F_{1,19} = 6.0$, $P = 0.02$, parental genotypes removed $R^2 = 0.21$, $F_{1,13} = 3.48$, $P = 0.08$; Spring migration: Full dataset $R^2 = 0.69$, $F_{1,28} = 61.88$, $P < 0.0001$, allopatric sites removed $R^2 = 0.54$, $F_{1,19} = 22.31$, $P = 0.0001$, parental genotypes removed $R^2 = 0.27$, $F_{1,13} = 4.72$, $P = 0.05$).

We evaluated the relationship between migratory behaviour and genetic background using ANOVAs as well, combining hybrids (i.e., those with scores between 1 and 4) into a single category and comparing the three taxonomic groups (i.e., coastal, inland and hybrid). We obtained similar results to the linear models described above and in the manuscript (Fall migration: Full dataset, $R^2 = 0.63$, $F = 23.39$, $P < 0.0001$, allopatric sites removed $R^2 = 0.23$, $F = 2.74$, $P = 0.09$; spring migration: Full dataset, $R^2 = 0.53$, $F = 15.14$, $P < 0.0001$, allopatric sites removed $R^2 = 0.34$, $F = 4.7$, $P = 0.02$; wintering grounds: Full dataset $R^2 = 0.70$, $F = 31.60$, $P < 0.0001$, allopatric sites removed $R^2 = 0.43$, $F = 6.85$, $P = 0.006$).

B.8 Measurement of traits for Cohen's D. Differences in migratory orientation, morphometrics, plumage colouration and song were quantified using Cohen's D. Differences in the first three traits were measured for allopatric birds (Delmore et al. 2012). Methods for summarizing migratory orientation are described the main text. Morphometric traits included wing, tarsus, tail and beak length, beak depth and width. Plumage traits included brightness, hue and saturation (red, green and UV chroma). We modified methods described in Reudink et al. (2009) for the plumage analysis. Briefly, we obtained six feathers from the mantel of each bird, mounted them on low reflectance black paper as they would naturally lie on a bird and used a UV-VIS-NIR light source with an Ocean Optics USB2000 spectrometer (Dunedin, FL, USA) for measurements. We took 10 readings for each bird and divided these data into 1 nm bins from 320 to 700 nm using CLR1.0.3 (Montgomerie 2008). We used principal component analyses to collapse each set of traits into one composite variable.

Ruegg et al. (2006b) measured a series of song traits in allopatric and hybrid populations of Swainson's thrushes. These authors used a discriminate function analysis to quantify differences between taxonomic groups; results can be found in Figure 2 of their paper. Differentiation was greatest along CV1, which accounted for 78% of the variation in song and on which song duration and frequency of the first one-fifth of the song loaded most strongly (0.362 and -0.189). We extracted the data for allopatric birds from this graph (n = 12 coastal, 17 inland).

Appendix C Supplementary material for Chapter 4

C.1 Additional methods for the assembly and annotation of draft reference genome. We trimmed raw sequencing files using a custom perl script to remove adaptor sequences, barcodes and bases with quality scores below 20 and used a reference based assembly method with the zebra finch genome as our template. All libraries were aligned to the zebra finch reference genome with Stampy-1.0.22 (Lunter and Goodson 2011), using a substitution rate of 5% to capture sequence variation expected due to divergence between the zebra finch and Swainson's thrush genomes. Stampy uses only paired reads in alignment and uses maximum likelihood to maximize mapping quality (pairing, correct insert size, orientation and errors). Aligned reads from all libraries were combined and used to build the consensus sequence for Swainson's thrush using *samtools* mpileup (Li et al. 2009) with the –E flag. This flag permits more sensitive BAQ calculations near sites with multiple polymorphisms and small indels, allowing thrush specific sequence to be represented, although there will still be a bias towards the zebra finch near large indels. *bcftools* view (with –cg flags) and custom perl scripts were used to convert output from mpileup to fastq. The zebra finch genome is commonly used in genomic studies of other birds (Lundberg et al. 2011; Ellegren et al. 2012; Parchman et al. 2013). Birds generally have a stable number of chromosomes, low rate of interchromosomal rearrangements, and high synteny (Dawson et al. 2007; Griffin et al. 2007; Backström et al. 2008; Stapley et al. 2008; Ellegren 2010).

We note that many genome studies (e.g., Ellegren et al. 2012) have used a different approach for their assemblies; they assemble their raw sequence denovo before aligning the resultant contigs and scaffolds to the zebra finch and/or a species-specific linkage map. This approach should reduce bias at a local scale and was attempted in our study but resulted in a less

complete assembly; the length of the final assembly was shorter (1,044,077,906 vs 1,233,099,904 bp), the N_{50} was smaller (54,324,215 vs 73,656,228 bp) and there were fewer genes (1,352 and 4,098 UCEs from the datasets described in the main text vs. 2,544 and 5,081). We also had far fewer re-sequencing reads align to the denovo assembly (35-37% vs. 67-71%). Given the goals of our study, to look at large scale patterns of divergence and identify candidate genes, we were willing to accept some bias on a localized scale for a more complete genome.

We masked the resultant genome using RepeatMasker4.0.3 (Smit et al. 1996) and annotated it used the MAKER pipeline (Cantarel et al. 2008). We used three *ab initio* gene predictors for this annotation: AUGUSTUS (Stanke et al. 2006), SNAP (Korf 2004) and Geneid (Blanco et al. 2007). AUGUSTUS and SNAP are available with MAKER; AUGUSTUS was run with the chicken training dataset; SNAP was trained with MAKER using ESTs and proteins for the zebra finch downloaded from Genbank (https://www.ncbi.nlm.nih.gov/genbank/) and UniProt (http://www.uniprot.org/), respectively. We collapsed ESTs into contigs using cap3 (Huang and Madan 1999) to remove redundancies. Geneid was run with a parameter file optimized for gene prediction in vertebrate and mammalian species. We also used gene predictions from the syntenic gene prediction tool SGP2. SGP2 combines *ab initio* gene prediction (Geneid) with tblastx searches between two or more genome sequences to provide both sensitive and specific gene predictions (Parra et al. 2003). The tblastx portion of SGP2 requires a reference genome to which the target genome is compared. We used the genome of the zebra finch (*T. guttata* ensembl assembly 3.2.4) as a reference to develop our parameter file for SGP2 because it seems to be at an appropriate evolutionary distance from target genome so that it is mostly the coding regions of the genes that seem to be significantly conserved between these two genomes. Moreover, we used an SGP2 parameter file optimized for predicting on the

chicken genome. We also included ESTs and proteins downloaded from Genbank and UniProt from both the zebra finch and chicken in the final run of MAKER and used InterProScan (Jones et al. 2014) to add functional information to the annotation, including protein domains and GO categories.

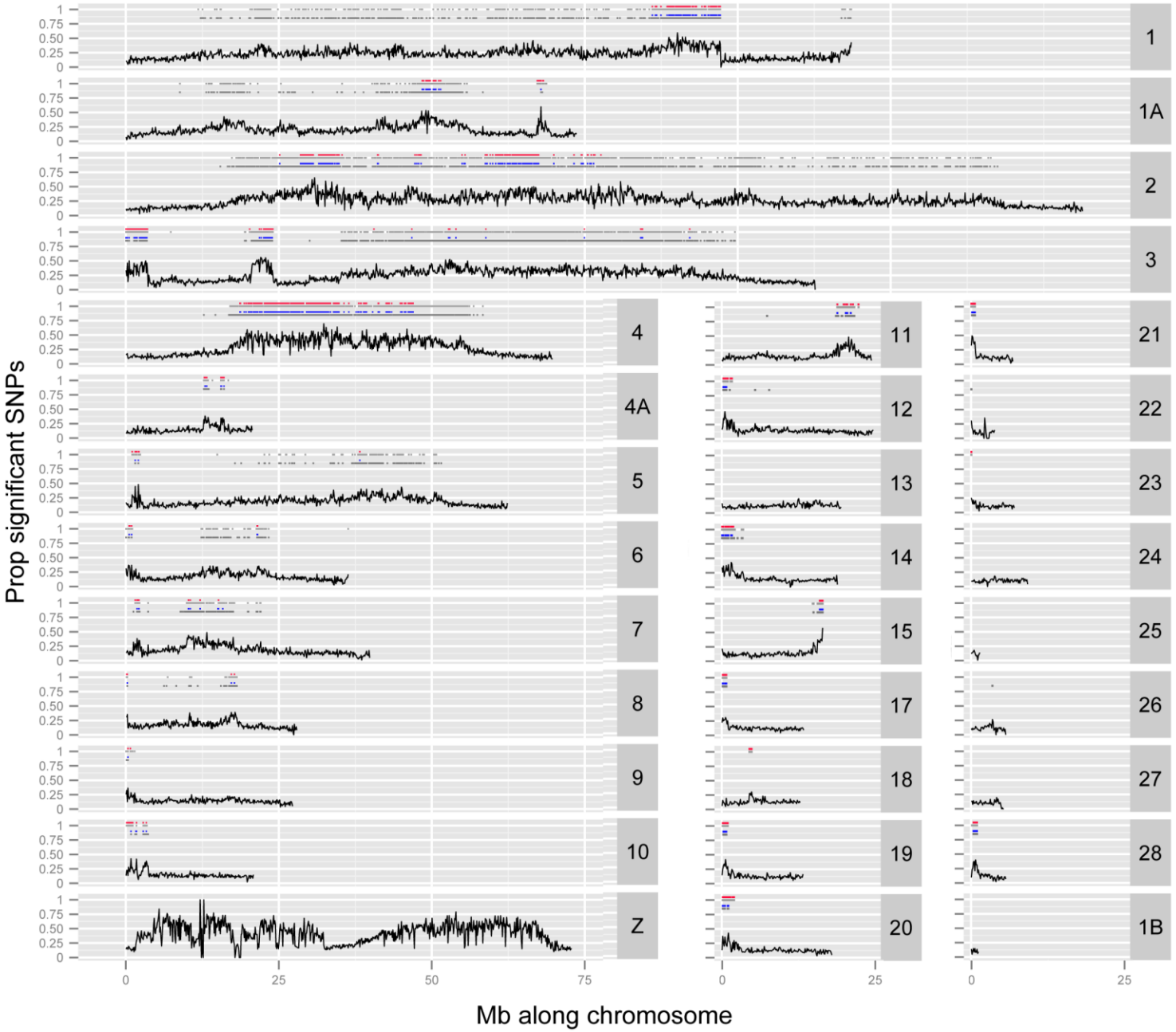
We inferred the location of centromeres in the genome following methods employed by (Ellegren et al. 2012), who assembled a reference genome for the collared flycatcher (personal communication, Reto Burri). We identified FISH probes from (Warren et al. 2010) on either side of centromeres in the zebra finch genome and used NCBI's blastn (Altschul et al. 1990) to find their location in our genome. We considered sequences between FISH probes as "centromeric regions". Information was only available for the larger, macrochromosomes (1-5,7,8, 1A and Z). We note that this method only gives us a rough approximation for the location of centromeres.

C.2 Number of reads sequenced for each library and those remaining after trimming and mapping steps. Data are shown for libraries used in (a) the draft reference genome, and (b) the re-sequencing analysis. Barcodes for each pool are as follows: GGCTAC for Kamloops, TAGCCT for Kelowna, CATCAG for Sunshine Coast and ACTTGA for Vancouver.

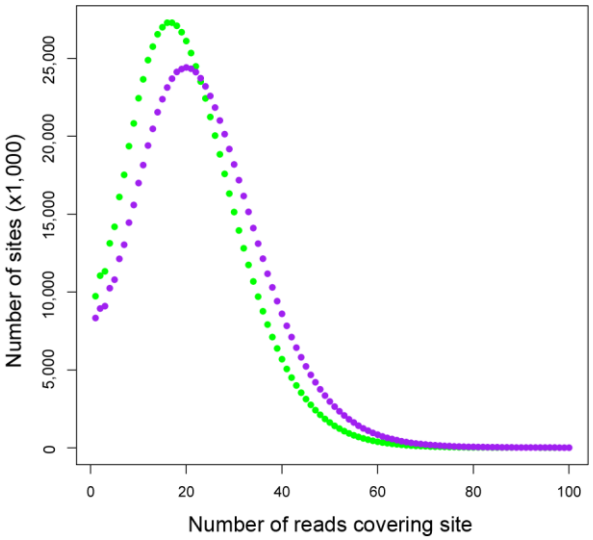
Library	**Raw reads**	**Trimmed reads**	**Mapped reads**
(a)			
Fragment	310,165,464	254,921,979	n/a
2kb mate pair	133,637,220	110,710,980	n/a
4kb mate pair	148,490,740	96,812,175	n/a
(b)			
Kamloops pool	180,809,234	170,875,944	119,697,268
Kelowna pool	218,344,764	208,419,466	148,824,508
Sunshine Coast pool	237,476,864	204,545,188	138,226,682
Vancouver pool	172,290,540	155,330,194	103,901,778

C.3 Additional methods for analyzing whole genome shotgun data from population samples. We trimmed raw sequencing files from each pool using the perl script described for assembling the reference genome and mapped reads to the reference and merged paired-end data in sam format with *bwa* mem (Li and Durbin 2009). We converted mapped reads from sam to bam using *samtools* view (q = 20; Li et al. 2009b) and merged files for each subspecies group using *samtools* merge. We identified and removed indels using GATK (RealignerTargetCreator and IndelRealigner, (DePristo et al. 2011). Default settings were used in all steps unless otherwise specified.

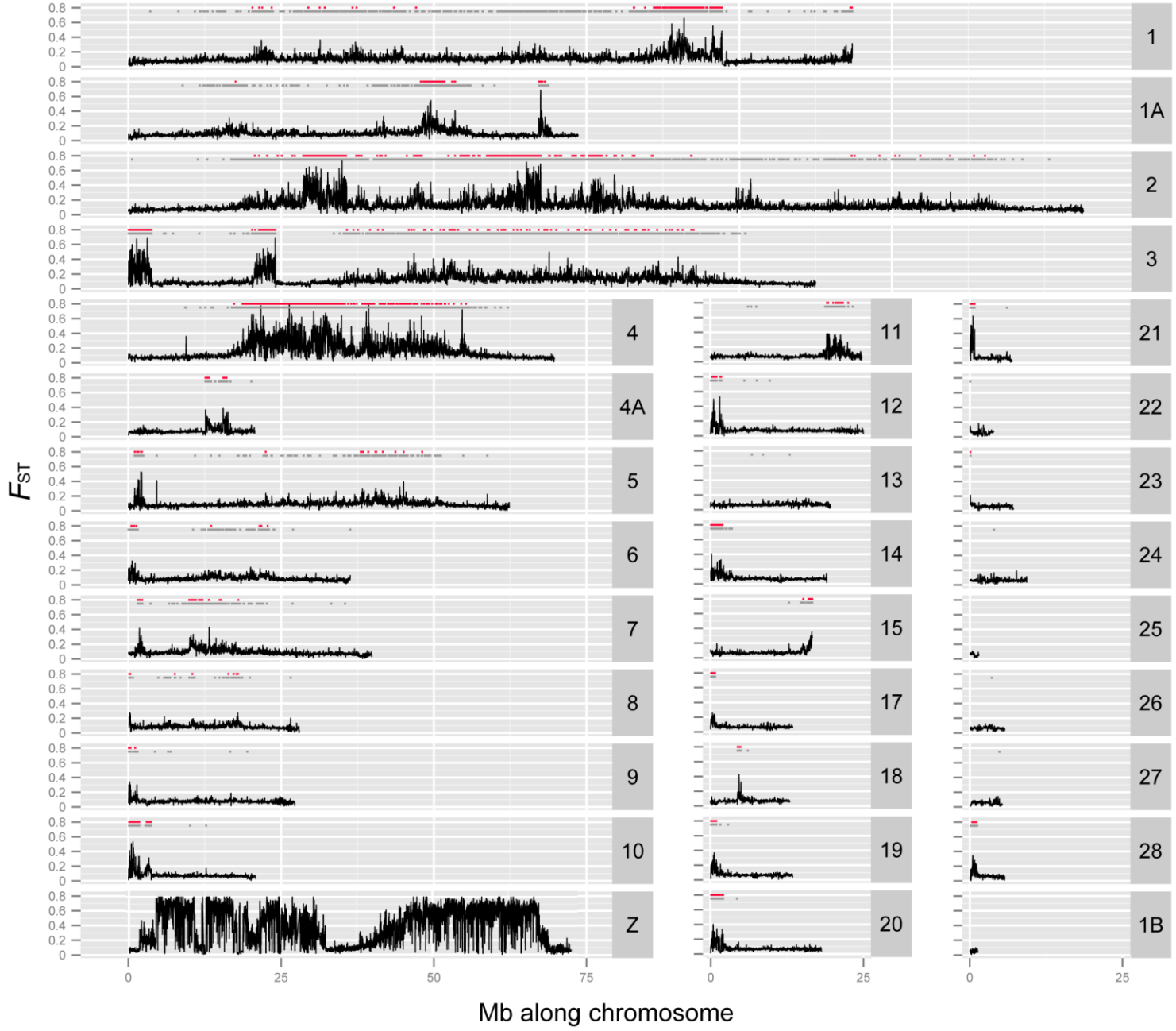
C.4 Proportion of significant SNPs in 100 kb, non-overlapping windows and based on a Fisher's Exact test (adjusted *P*-value of 0.01, method = fdr). Dark gray lines show outlier windows identified using a resampling test, comparing the proportion of significant SNPs in a window to the number in 1000 resamples from the rest of the genome. Blue lines show a subset of the latter windows in the bottom 5% of Tajima's D values. Light gray and red lines show outlier windows from Figure 4.1, identified using F_{ST} and Tajima's *D*.


Prop significant SNPs
Mb along chromosome
1
1A
2
3
4
4A
5
6
7
8
9
10
Z
11
12
13
14
15
17
18
19
20
21
22
23
24
25
26
27
28
1B

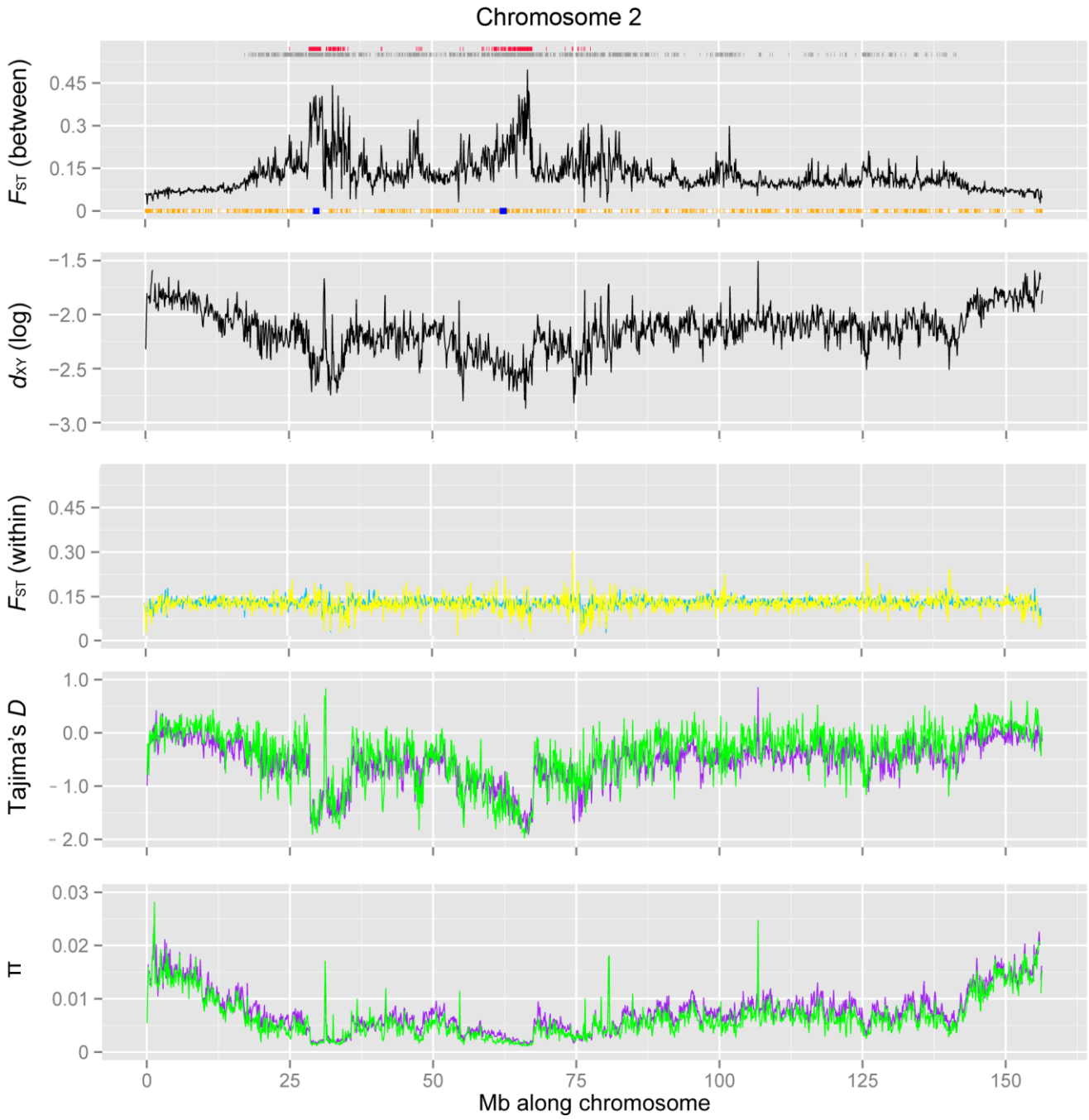
C.5 Coverage distribution of coastal (green) and inland (purple) pools.

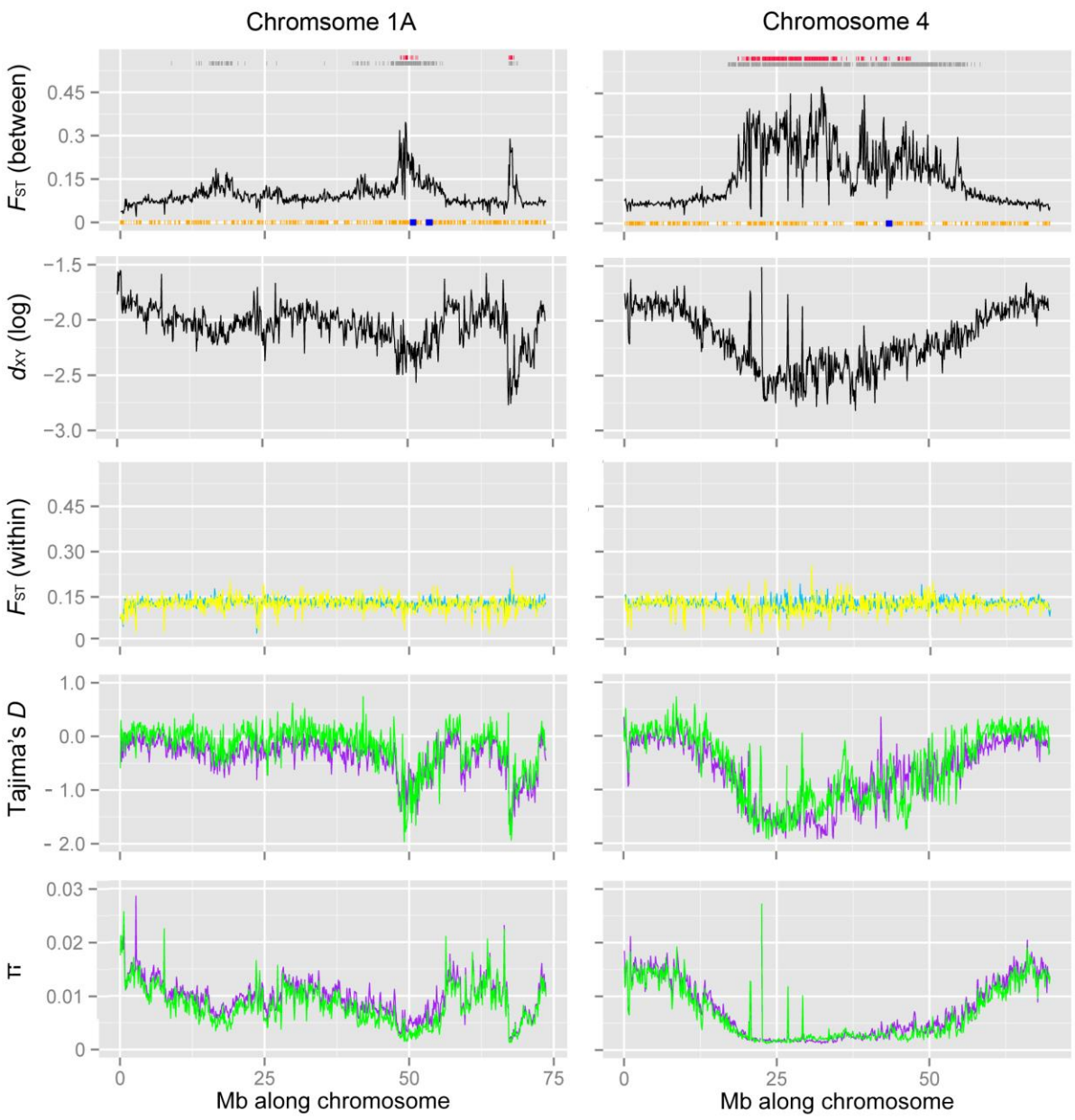


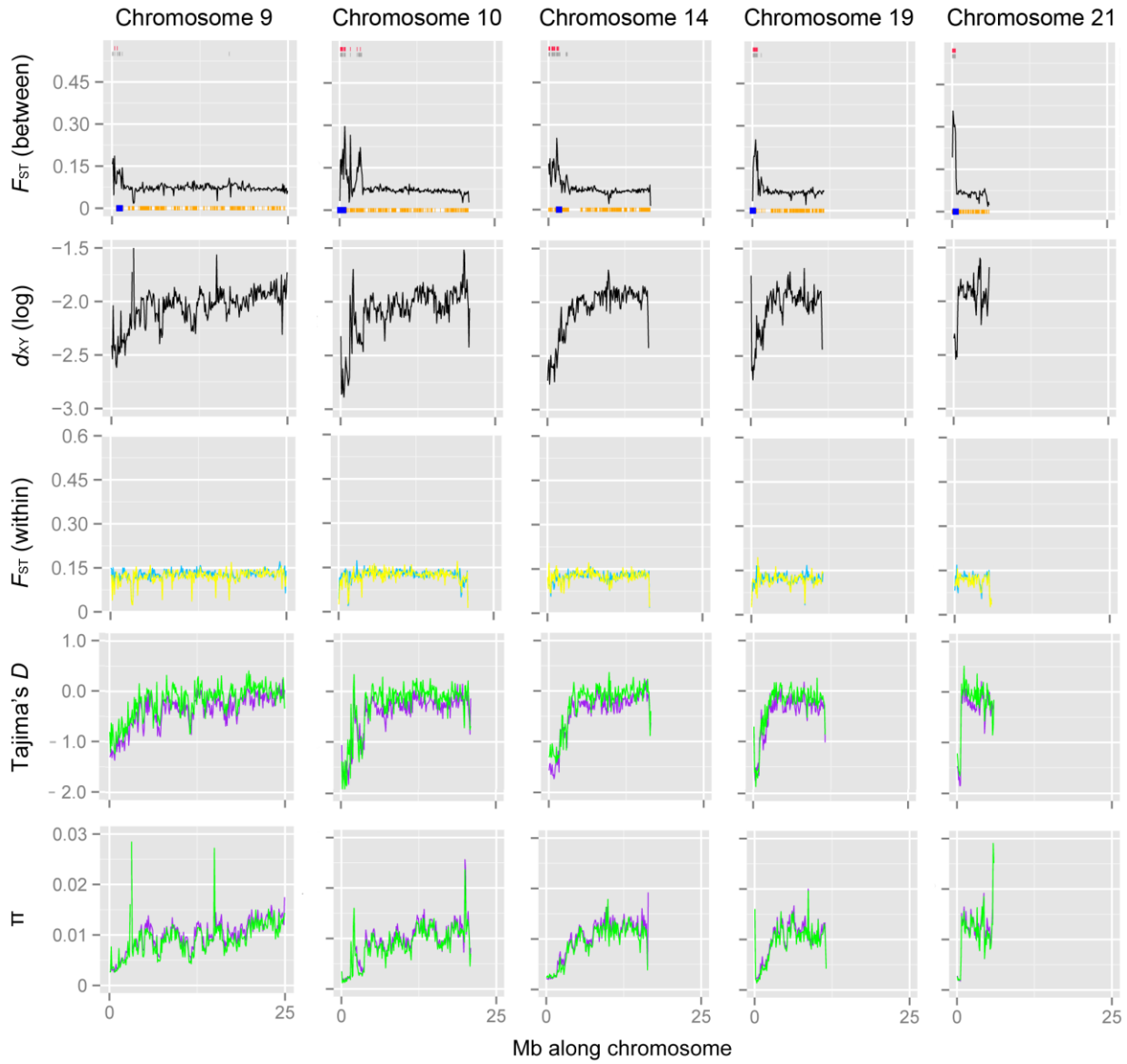
C.6 F_{ST} estimated in 20 kb, non-overlapping windows and between populations of coastal and inland subspecies groups of Swainson's thrushes. Gray lines at the top of each panel show F_{ST} outlier windows identified using a HMM; red lines identify F_{ST} outlier windows that were also in the bottom 5% of Tajima's *D* values. F_{ST} was summarized for the Z chromosome but this chromosome was not included in outlier analyses (since it is on average so highly divergent between the subspecies groups and has a different mode of inheritance than the autosomes).


F_{ST}
0.8
0.6
0.4
0.2
0
1
1A
2
3
4
4A
5
6
7
8
9
10
Z
11
12
13
14
15
17
18
19
20
21
22
23
24
25
26
27
28
1B
0
25
50
75
Mb along chromosome

C.7 Population genomic estimates for autosomes that had candidate genes in one or more outlier windows. Estimates are from non-overlapping windows of 100 kb. The first panel shows data from Figure 4.2 in more detail for these chromosomes, including estimates of F_{ST} between the subspecies, outliers identified using both F_{ST} and Tajima's D (red), outliers identified using F_{ST} (gray), all of the genes located on this chromosome (orange) and candidate genes found in outlier windows (blue; Table 4.2 and Appendix C.8). The remaining panels show d_{XY}, F_{ST} estimated between populations within subspecies groups (estimates between coastal populations in blue, between inland populations in yellow), Tajima's D (estimates for inland populations in purple, for coastal populations in green) and nucleotide diversity (estimates coloured as in panel d).


Chromosome 2
F_{ST} (between)
0.45
0.3
0.15
0
d_{XY} (log)
−1.5
−2.0
−2.5
−3.0
F_{ST} (within)
0.45
0.30
0.15
0
Tajima's D
1.0
0.0
- 1.0
- 2.0
π
0.03
0.02
0.01
0
0
25
50
75
100
125
150
Mb along chromosome


Chromsome 1A
Chromosome 4
F_{ST} (between)
0.45
0.3
0.15
0
d_{XY} (log)
−1.5
−2.0
−2.5
−3.0
F_{ST} (within)
0.45
0.30
0.15
0
Tajima's D
1.0
0.0
- 1.0
- 2.0
π
0.03
0.02
0.01
0
0
25
50
75
Mb along chromosome
0
25
50
Mb along chromosome


Chromosome 9
Chromosome 10
Chromosome 14
Chromosome 19
Chromosome 21
F_{ST} (between)
0.45
0.30
0.15
0
d_{XY} (log)
−1.5
−2.0
−2.5
−3.0
F_{ST} (within)
0.6
0.45
0.30
0.15
0
Tajima's D
1.0
0.0
- 1.0
- 2.0
π
0.03
0.02
0.01
0
0
25
Mb along chromosome

C.8 List of candidate and functionally related genes generated in the present study using a key word search in Ensembl. Information on the gene's location in our annotation (Chrom) is provided, along with whether the gene was present in outlier windows of 1000, 100 and 20 kb. The proportion of non-synonymous (NS) SNPs in each gene found in an outlier window is also provided.

Family Name	Gene	Chr	1000	100	20	NS SNPs
period genes	PER2	9	N	N	**Y**	3/8
	PER3	21	N	N	N	
crytochromes	CRY1	1A	N	N	**Y**	0/2
	CRY2	5	N	N	N	
F-box proteins,	FBXL3	1	N	N	N	
leucine-rich repeat class	FBXL21	13	N	N	N	
	FBXL18	14	N	N	N	
	FBXL2	2	**Y**	**Y**	**Y**	1/2
	FBXL15	6	N	N	N	
	FBXL22	10	**Y**	**Y**	**Y**	1/1
	FBXL16	14	**Y**	N	**Y**	1/4
	FBXL4	3	N	N	N	
	FBXL5	4	N	N	N	
	FBXL7	2	N	N	N	
	FBXL20	-	n/a	n/a	n/a	
Basic helix-loop-helix proteins	ATOH7	-	n/a	n/a	n/a	
	BHLHE40	12	N	N	N	
	NPAS2	1	N	N	N	
	BHLHE41	1A	N	N	N	
	CLOCK	4	N	**Y**	**Y**	1/3
	TCF25	11	N	N	N	
	TCF15	20	N	N	N	
	TCFL5	20	N	N	N	
	KIAA2018	1	**Y**	**Y**	**Y**	4/7
	NPAS3	5	N	N	N	
	HIF1A	5	N	N	N	
	ARNTL2	1A	N	N	N	
	EPAS1	3	N	N	N	

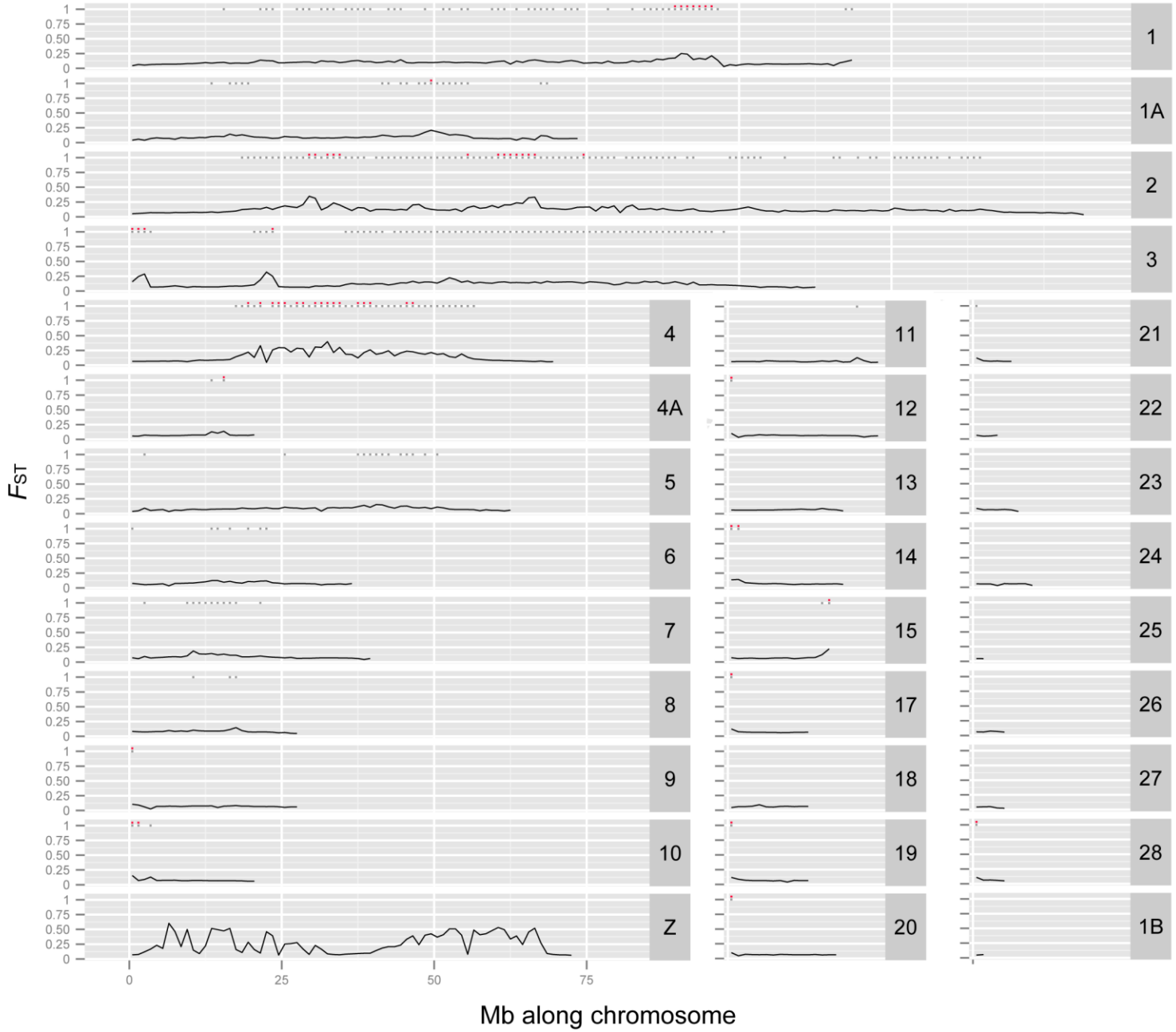
Family Name	Gene	Chr	1000	100	20	NS SNPs
	MLXIPL	19	**Y**	**Y**	**Y**	2/5
	FERD3L	-	n/a	n/a	n/a	
	OLIG3	-	n/a	n/a	n/a	
	NEUROG1	-	n/a	n/a	n/a	
	TWIST1	-	n/a	n/a	n/a	
	TWIST2	-	n/a	n/a	n/a	
Dopamine receptors	DRD2	24	N	N	N	
	DRD3	1	**Y**	**Y**	**Y**	3/4
	DRD4	5	N	N	N	
	DRD1	13	N	N	N	
	DRD5	-	n/a	n/a	n/a	
Neuronal acetylcholine receptors	CHRNB2	25	N	N	N	
	CHRNB3	Z	n/a	n/a	n/a	
	CHRNA9	4	N	N	N	
	CHRNA3	10	**Y**	**Y**	**Y**	3/4
	CHRNA4	20	N	N	N	
	CHRNA5	10	**Y**	**Y**	N	2/3
casein kinase I	CSNK1D	-	n/a	n/a	n/a	
	CSNK1E	1A	N	**Y**	**Y**	2/3
	CSNK1A1	13	N	N	N	
	CSNK1G1	10	N	N	N	
Sine oculis homeobox	SIX3	-	n/a	n/a	n/a	
transcription factors	SIX2	3	N	N	N	
	SIX4	5	N	N	N	
	SIX6	5	N	N	N	
Opsins	OPN4	6	N	N	N	
	OPN3	3	N	N	N	
	OPN5	3	N	N	N	
	OPN1LW	19	N	N	N	
	RHO-2	12	N	N	N	
	RHO-1	-	n/a	n/a	n/a	
Serotonin receptors	HTR7	4	N	N	N	
	HTR1D	23	N	N	N	
	HTR2A	1	N	N	N	
	HTR3A	24	N	N	N	
	HTR4	13	N	N	N	
	HTR1E	-	n/a	n/a	n/a	
	HTR1F	-	n/a	n/a	n/a	
	HTR1B	-	n/a	n/a	n/a	

Family Name	Gene	Chr	1000	100	20	NS SNPs
	HTR2C	-	n/a	n/a	n/a	
	HTR2B	-	n/a	n/a	n/a	
	HTR5A	-	n/a	n/a	n/a	
Solute carrier family 4,	SLC4A11	4	N	N	N	
bicarbonate transporters	SLC4A2	2	N	N	N	
	SLC4A3	7	N	N	N	
	SLC4A4	19	N	N	N	
	SLC4A9	13	N	N	N	
	SLC4A10	7	N	N	N	
	SLC4A1	27	N	N	N	
	SLC4A7	-	n/a	n/a	n/a	
dual-specificity tyrosine	DYRK1A	1	N	N	N	
phosphorylation-	DYRK4	1A	N	N	N	
regulated kinases	DYRK3	26	N	N	N	
	DYRK2	-	n/a	n/a	n/a	
N-Cor repressor complex	NCOR1	19	N	N	N	
	TBL1XR1	9	N	N	N	
	TBL1X	1	N	N	N	
	NCOR2	-	n/a	n/a	n/a	
	HDAC3	-	n/a	n/a	n/a	
BK family of calcium activated	KCNMA1	6	N	N	N	
potassium channels	KCNMB2	9	N	N	N	
	KCNMB4	1A	N	N	N	
	KCNMB1	-	n/a	n/a	n/a	
Glycogen synthase kinase-3	GSK3B	-	n/a	n/a	n/a	
	GSK3A	-	n/a	n/a	n/a	
Hypocretin receptor	HCRTR2	3	N	N	N	
Inhibitor of DNA binding proteins	ID2	-	n/a	n/a	n/a	
	ID3	23	N	N	N	
	ID4	-	n/a	n/a	n/a	
Senataxin	SETX	17	N	N	N	
Prokineticin receptors	PROK2	12	N	N	N	
	PROKR1	-	n/a	n/a	n/a	

The following genes had more than 20 related genes and were not included

Dead-box proteins	DDX5	18
Serine/threonine protein kinase	PRKG2	4
BTB/POZ domain-containing prot	BTBD9	3

C.9 F_{ST} estimated in 1000 kb, non-overlapping windows estimated between populations of inland and coastal subspecies groups of Swainson's thrushes. Gray lines at the top of each panel identify F_{ST} outlier windows identified using a HMM; red lines identify F_{ST} outlier windows that were also in the bottom 5% of Tajima's *D* estimates. F_{ST} was summarized for the Z chromosome but this chromosome was not included in outlier analyses.

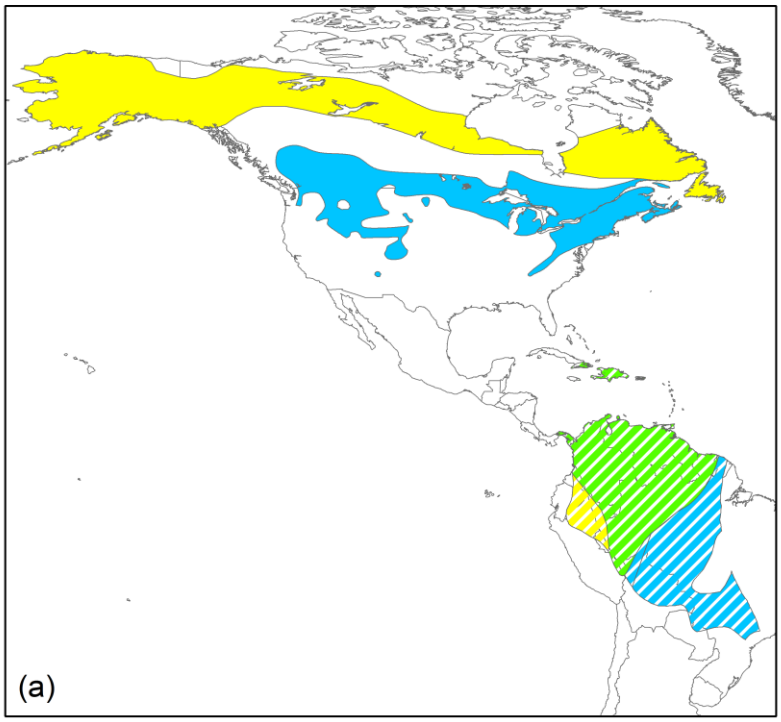

F_{ST}
1
0.75
0.50
0.25
0
1A
2
3
4
4A
5
6
7
8
9
10
Z
11
12
13
14
15
17
18
19
20
21
22
23
24
25
26
27
28
1B
0
25
50
75
Mb along chromosome

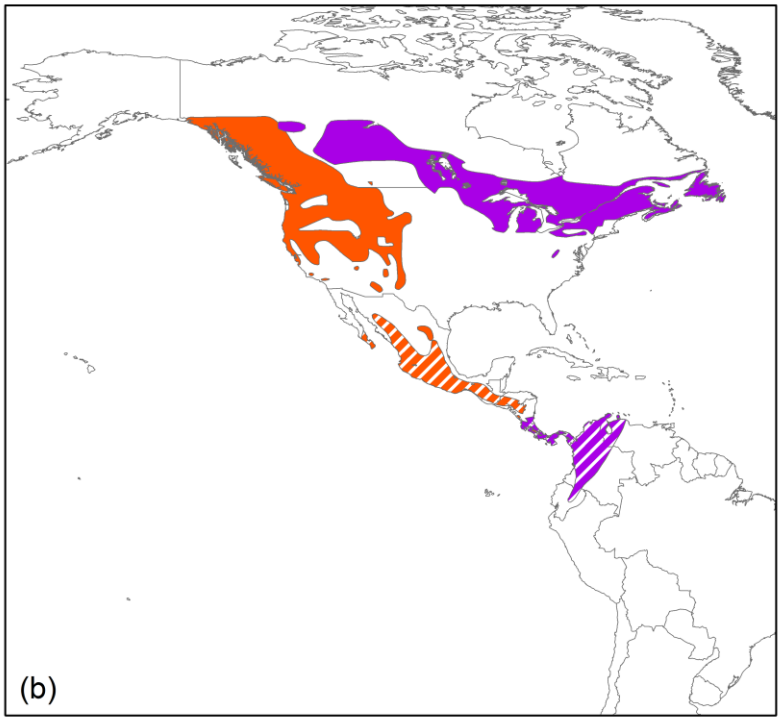
C.10 Gene ontology (GO) categories enriched in outlier windows and the genes associated with these categories.

GO	Ontology	Description	Name in thrush annotation	Interpro name
0043515	kinetochore binding	molecular function	maker-2-snap-gene-149.31	CLASP2
			maker-2-snap-gene-149.32	Uncharacterized
			maker-4-snap-gene-108.20	Uncharacterized
0051225	spindle assembly	biological process	maker-4-augustus-gene-142.14	MAP9
			maker-1-augustus-gene-451.39	CHD4
			maker-4-snap-gene-132.49	TNKS

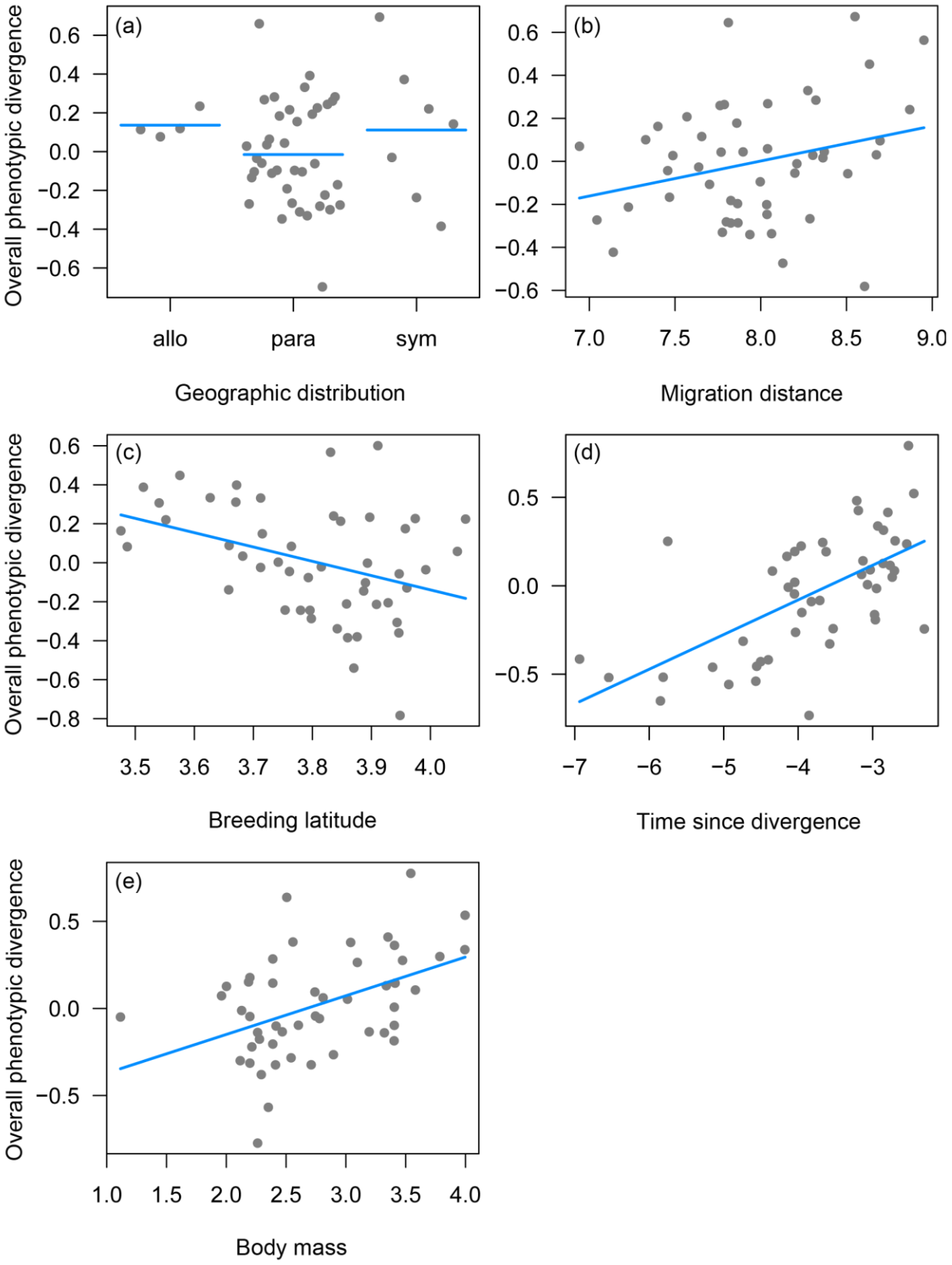
Appendix D Supplementary material for Chapter 5

D.1 Breeding and wintering ranges (hatched) for two species pairs included in analysis, including (a) a pair that does not form a migratory divide (*Catharus minimus* [Gray-cheeked thrush, yellow] and *Catharus fuscescens* [Veery, blue]) and (b) a pair that does form a migratory divide (*Geothlypis tolmiei* [MacGillivray's warbler, orange] and *Geothlypis philadelphia* [Mourning warbler, purple]). Green shows the overlap between ranges in (a).


(a)


(b)

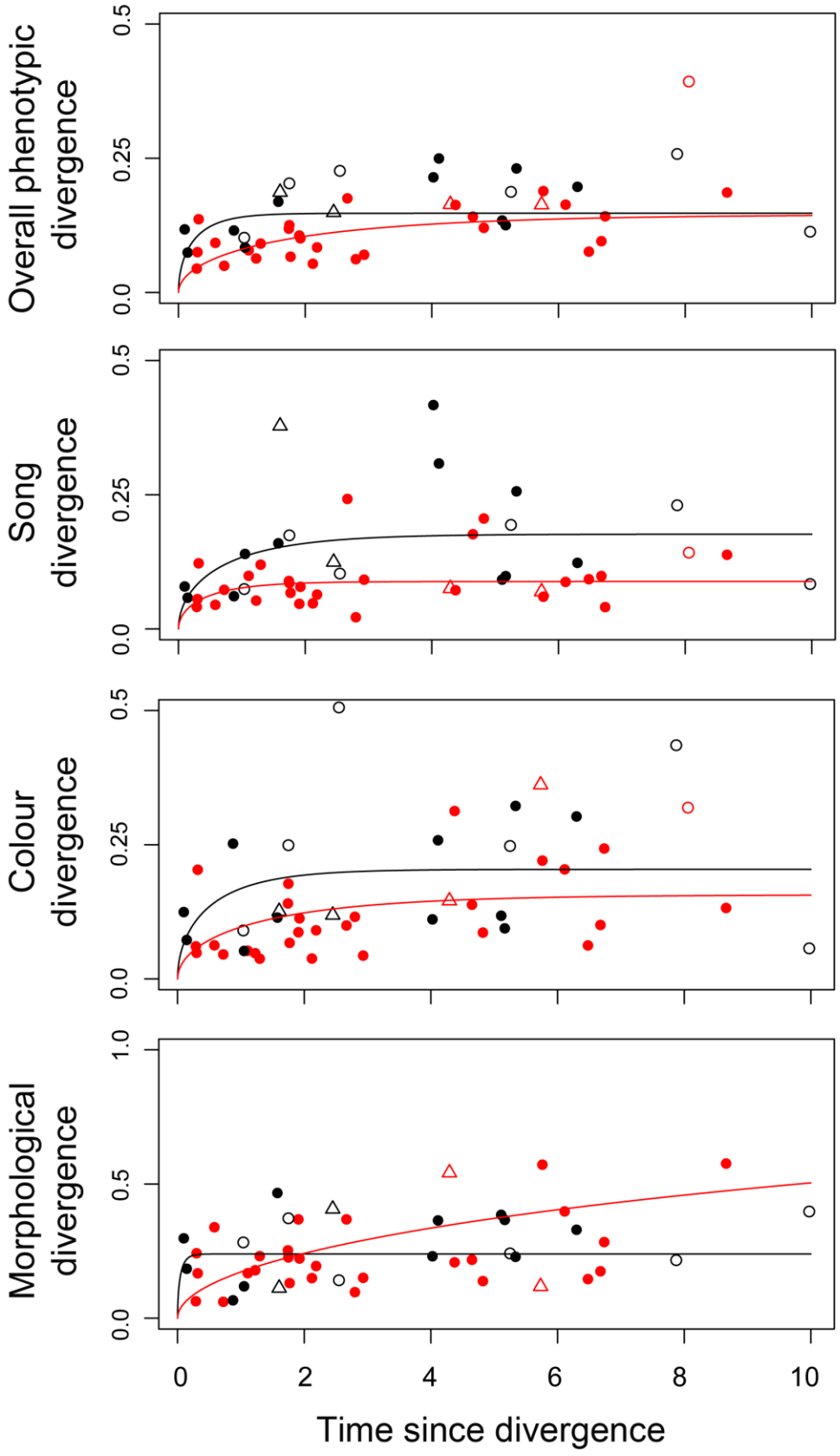
D.2 Modelled relationship between overall phenotypic divergence and five additional predictor variables included in global linear mixed effect models. Migration category was also included (relationship shown in Figure 5.1) and a random variable of genus nested within family to control for phylogenetic relationships. Significant predictors in the final model included migration category, time since divergence and body mass. The effect each predictor is plotted with these additional variables held at their medians.


(a)
Overall phenotypic divergence
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
allo
para
sym
Geographic distribution
(b)
7.0
7.5
8.0
8.5
9.0
Migration distance
(c)
−0.8
3.5
3.6
3.7
3.8
3.9
4.0
Breeding latitude
(d)
0.5
−0.5
−7
−6
−5
−4
−3
Time since divergence
(e)
1.0
1.5
2.0
2.5
3.0
3.5
4.0
Body mass

D.3 Additional results from Ornstein–Ulhenbeck models of evolutionary change in phenotypic traits. Models included separate rates for each migratory candidate and were run for each trait set (i.e., combined and separate). One set of models included geographic distribution (geography; continuous variable from 0 [allopatric] to 2 [sympatric]) and the other did not (results also shown in Table 2). Log-likelihood values for each model are shown and are all within 1.92 units suggesting there is no significant difference between them. Values for the rate (β) of change in phenotypic divergence with geography are shown along with confidence intervals (CI; lower, upper) for models that included this variable as a covariate. All CI overlap 0 suggesting rates are non-significant.

		No divide		Divide	
	Log-likelihood	β	CI	β	CI
a) Overall phenotypic divergence					
no geography	25.63		n/a		
geography	26.44	0.0010	-0.020, 0.044	0.0010	-0.10, 0.0078
b) Song divergence					
no geography	31.73		n/a		
geography	32.97	-0.010	-0.022, 0.009	0.0042	0, 0.061
c) Colour divergence					
no geography	16.43		n/a		
geography	17.51	0.0019	-0.0080, 0.052	0.0023	-0.10, 0.10
d) Morphological divergence					
no geography	21.65		n/a		
geography	23.42	-0.010	-0.10, 0.58	0.0026	-0.067, 0.012

D.4 Relationship between phenotypic divergence and time since divergence including geographic distribution. Results are shown for all traits combined (a) and each set of traits separately (b to d). Time since divergence was measured using genetic distance; phenotypic divergence was measured using average Hedges' *g*. Curves show results from OU models of evolutionary change, fitting separate models for taxa that do (red line and points) and do not (black line and points) form migratory divides and including geographic distribution as a predictor. Allopatric pairs are shown in open triangles, parapatric pairs with closed circles and sympatric pairs with open circles.


Overall phenotypic divergence
Song divergence
Colour divergence
Morphological divergence
0.0
0.25
0.5
1.0
0
2
4
6
8
10
Time since divergence