COMBINING FIELD-BASED BEHAVIORAL EXPERIMENTS WITH GENOMICS TO UNDERSTAND RAPID AVIAN SPECIATION

by

SHEELA PHANSALKAR TURBEK

B.A., Bowdoin College, 2013

A thesis submitted to the

Faculty of the Graduate School of

University of Colorado in partial fulfillment

of the requirements of the degree of

Doctor of Philosophy

Department of Ecology and Evolutionary Biology

2021

Committee Members:

Scott Taylor

Leonardo Campagna

Michael Breed

Nancy Emery

Andrew McAdam

Turbek, Sheela Phansalkar (PhD., Ecology and Evolutionary Biology)

Combining field-based behavioral experiments with genomics to understand rapid avian speciation

Thesis directed by Dr. Scott A. Taylor

ABSTRACT

Behavioral and phenotypic traits play an important role early in speciation by influencing 1) where and when individuals come into contact and 2) whether closely related organisms are recognized as potential mates. Yet, despite the key role of pre-mating isolation in the evolution of biodiversity, our understanding of the specific mechanisms by which phenotypes and behavioral processes contribute to the generation of reproductive isolation during incipient speciation remains limited. My dissertation research examines the ways in which migratory behavior and sexual signals influence gene flow in two avian radiations in the early stages of divergence: southern capuchino seedeaters (*Sporophila*) and the barn swallow species complex (*Hirundo rustica*). First, I present a conceptual chapter that synthesizes current literature and organizes hypothesis testing about the ways in which behavioral and phenotypic traits, specifically migratory strategy, may mediate patterns of gene flow early in the speciation process. Then, I examine the importance of divergent migratory behavior in the evolution of two subspecies of barn swallow (*H. r. rustica* and *H. r. gutturalis*) that form a hybrid zone in Gansu Province, China. The subspecies exhibit a striking migratory divide that spans two continents and is closely associated with genomic differentiation across the hybrid zone, suggesting that assortative mating by timing of arrival and/or selection against hybrids that inherit intermediate migratory traits may limit interbreeding between the subspecies. My fourth chapter analyzes the

genomic and behavioral bases of pre-mating isolation between two species of capuchino seedeaters (*S. hypoxantha* and *S. iberaensis*) that co-occur during the breeding season in Iberá National Park, Argentina. Though the species lack obvious ecological barriers to reproduction, I document behaviorally-mediated species recognition and strong assortative mating associated with genomic regions underlying male plumage patterning. Finally, I generate fine-scale recombination maps for capuchino seedeaters to examine the role that variation in recombination rate has played in generating phenotypic diversity and peaks of genomic differentiation early in the speciation process. By combining fine-scale behavioral analyses with phenotype data and high-throughput genomic sequencing, these chapters investigate the traits underlying reproductive isolation and their implication for speciation in recent avian radiations.

ACKNOWLEDGEMENTS

This research would not have been possible without the unwavering support and dedicated mentorship of my advisor Dr. Scott Taylor and collaborator Dr. Leonardo Campagna. I thank them both for believing in this work and encouraging me to aim as high as possible. The Taylor Lab has been a wonderful environment to work in, and I am grateful to my current and former lab-mates for their constructive feedback and guidance throughout my time in graduate school. Thanks to my committee - Dr. Mike Breed, Dr. Nancy Emery, and Dr. Andrew McAdam - for their thoughtful comments on my dissertation research. This work was funded by a National Science Foundation Graduate Research Fellowship, the Bernice Udick Graduate School Dissertation Completion Fellowship, the University of Colorado Boulder EBIO Summer Fellowship, and grants to SPT from the American Genetics Association, the National Geographic Society, the American Philosophical Society, the University of Colorado Boulder, the American Society of Naturalists, the Wilson Ornithological Society, the American Ornithological Society, the British Ornithologists' Union, the Society of Systematic Biologists, the Society for Integrative and Comparative Biology, the American Museum of Natural History, the Cooper Ornithological Society, and the Animal Behavior Society.

Chapter II. I thank my co-authors Elizabeth Scordato and Rebecca Safran for their assistance with this opinion piece, Scott Taylor, Amanda Hund, and Kathryn Grabenstein for their helpful suggestions regarding the pre-submission inquiry, Kira Delmore for her valuable input on an early version of the case study table, and J Albert Uy and Darren Irwin, as well as three

anonymous reviewers, for their thoughtful comments on the manuscript. RJS and ESCS were supported by a National Science Foundation CAREER grant (DEB-CAREER 1149942 to RJS).

Chapter III. Thank you to my co-authors Drew Schield, Elizabeth Scordato, Andrea Contina, Xin-Wei Da, Yang Liu, Yu Liu, Emilio Pagani-Núñez, Qing-Miao Ren, Chris Smith, Craig Stricker, Michael Wunder, David Zonana, and Rebecca Safran for assisting with field work, helping with the genomic and stable isotope analyses, and providing useful feedback on the manuscript. I would additionally like to thank Lanzhou University (particularly Dr. Bo Du), Sun Yat-sen University, the State Darwin Museum in Moscow, Hainan Normal University, National University of Mongolia, the Mongolian Ornithological Society, and the Japan Bird Research Association for sponsoring our research. I thank Zach Gompert for his assistance with the ddRAD lane effect issue. This work was supported by a National Science Foundation Postdoctoral Fellowship to DRS (DBI-1906188), a National Geographic Society Committee on Research and Exploration grant to ESCS, and National Science Foundation grants to RJS (DEB-CAREER 1149942, IOS 1856266) and MW (DBI-1564947).

Chapter IV. I thank my co-authors Melanie Browne, Adrián S. Di Giacomo, Cecilia Kopuchian, Wesley Hochachka, Cecilia Estalles, Darío Lijtmaer, Pablo Tubaro, Luís Fábio Silveira, Irby Lovette, Rebecca J. Safran, Scott Taylor, and Leonardo Campagna for their help with project design, field work, obtaining and analyzing genomic data, and analyzing eBird data. I would additionally like to thank Constanza Pasian, Bárbara Rodríguez-Quintana, and Bruna Rodrigues do Amaral for field assistance, Ethan Turbek for generating the mounts, and feedback from the Taylor, Safran, and Lovette lab groups, as well as three anonymous reviewers. Adrián Di

Giacomo, Cecilia Kopuchian, Melanie Browne, Sofia Zalazar, Constanza Pasian, and Florencia Pucheta welcomed me to Argentina with open arms, and I am incredibly grateful to them for all of their logistical support, including driving me to and from the airport, providing me with housing, and showing me around Corrientes on countless occasions.

Chapter V. Thanks to Georgy Semenov, Erik Enbody, Leonardo Campagna, and Scott Taylor, who assisted with this project, as well as Per Alström and Luís Fábio Silveira with the Museu de Zoologia da Universidade de São Paulo for providing access to the wagtail and capuchino outgroup samples, respectively. In addition, I acknowledge support from the National Genomics Infrastructure in Genomics Production Stockholm funded by Science for Life Laboratory and SNIC/Uppsala Multidisciplinary Center for Advanced Computational Science for assistance with parallel sequencing and access to the UPPMAX computational infrastructure.

CONTENTS

LIST OF TABLES

TABLE

LIST OF FIGURES

FIGURE

CHAPTER 1

INTRODUCTION

Identifying the mechanisms that promote divergence along each stage of the speciation continuum, from interbreeding populations to fully reproductively isolated taxa, is an enduring challenge of evolutionary biology. Initial investigations into the speciation process largely focused on the contribution of geographic context to patterns of genetic exchange (Mayr 1963; Coyne and Orr 2004). Geographic isolation promotes genetic and phenotypic divergence by limiting gene flow and allowing the buildup of isolating mechanisms between closely related populations. However, growing evidence indicates that genomic differentiation can proceed despite ongoing genetic exchange between divergent lineages (Nosil 2008; Matute 2010; Pinho and Hey 2010; Butlin et al. 2013; Martin et al. 2013; Feder et al. 2018). In some cases, gene flow can even introduce novel and beneficial genetic variation into a population that selection can then act upon (Seehausen 2004; Marques et al. 2019). In areas of geographic overlap, the buildup and/or maintenance of reproductive isolation occurs via divergent natural and sexual selection (Coyne and Orr 2004).

A well-supported mechanism of speciation is ecological speciation, in which reproductive isolation evolves between closely related organisms as a result of adaptation to different environments (Schluter 2009). Ecologically-based divergent selection can arise through a variety of means, including resource competition, pollinator attraction, and predator avoidance, and is thought to drive speciation or maintain species boundaries in numerous systems [*e.g.*, insects adapted to different host plants (Forbes et al. 2017), plants that exhibit variation in flowering time (Taylor and Friesen 2017), and fish that mate assortatively by body size (Ryan et

al. 2016)]. Speciation by sexual selection, on the other hand, occurs when parallel changes in mate preferences and sexual signals within populations lead to behavioral isolation between populations (West-Eberhard 1983; Panhuis et al. 2001; Ritchie 2007; Servedio and Boughman 2017). Sexual selection results from within-population variation in the ability of individuals to acquire mates, and is often caused by male-male competition or female choice (Andersson 1994; Servedio and Boughman 2017). Recent studies have therefore attempted to expand the definition of speciation by sexual selection to consider the role of male-male competition in divergence and the maintenance of reproductive isolation (Keagy et al. 2016; Tinghitella et al. 2017). Though less well understood than ecological speciation, speciation by sexual selection is generally invoked to explain divergence in systems that inhabit similar environments or differ solely in secondary sexual characters (Tinghitella et al. 2017). In reality, natural and sexual selection operate simultaneously within populations and likely interact during the speciation process (Maan and Seehausen 2011; Safran et al. 2013; Scordato and Safran 2014). Both processes ultimately lead to the generation of reproductive isolation, or the absence or reduction of gene flow between closely related groups either prior to or following mating.

Pre-mating isolation describes the processes that lead to assortative mating, or the reproductive pairing of similar individuals, and restrict gene flow prior to fertilization (Jiang et al. 2013). While assortative mating is hypothesized to play a critical role in speciation by forming or maintaining barriers to genetic exchange, the specific behavioral processes and traits underlying assortative interactions and their relative roles in the evolution of reproductive isolation are often poorly understood (Kopp et al. 2017). Pre-mating isolation can arise through a variety of means, including spatial confinement to different habitats, temporal segregation of breeding, or divergence in mate preferences or traits used in species recognition between closely

related populations (Coyne and Orr 2004; Kopp et al. 2017). Habitat isolation occurs when closely related groups develop divergent habitat preferences and mating occurs close to preferred habitat, thereby reducing the probability of heterospecific encounters (Rundle and Nosil 2005; Tonnis et al. 2005; Rolshausen et al. 2013; Xue et al. 2014). Likewise, temporal isolation is generated when closely related populations mate at different times of the year due to divergent developmental schedules or life history strategies, again reducing the likelihood of encounters with receptive heterospecifics (Friesen et al. 2007; Taylor and Friesen 2017). Both habitat isolation and temporal isolation are likely common during ecological speciation, as adaptation to divergent environments can affect both habitat preferences and timing of breeding (Rundle and Nosil 2005). In contrast, behavioral isolation occurs when individuals from different populations evolve divergent mate preferences or fail to recognize one another as potential mates (Rundle and Nosil 2005; Mendelson et al. 2017).

In cases where assortative mating is not sufficiently strong to prevent homogenization, post-mating mechanisms, such as genetic incompatibilities or ecological and/or sexual selection against hybrids, can further limit genetic exchange between divergent lineages (Coyne and Orr 2004; Matute et al. 2014). Genetic incompatibilities are produced when alleles at different genetic loci that do not function well together are combined in hybrids, causing hybrid inviability or sterility (Matute et al. 2014). Hybrids may also experience reduced fitness if they possess intermediate phenotypes that prevent them from using parental niches or securing mates (Svedin et al. 2008). Post-mating reproductive isolation can in turn select for strengthened pre-mating isolation through the process known as reinforcement (Howard 1993; Coyne and Orr 2004). During reinforcement, selection for increased mate discrimination to avoid maladaptive

hybridization is expected to yield greater divergence in traits associated with mate choice (Hoskin et al. 2005; Higgie and Blows 2007; Pfennig and Rice 2014).

While pre-mating isolation is often cited as a more effective barrier to gene flow than post-mating isolation due to the sequential nature of isolating barriers, the relative significance of these processes extremely early in speciation is unclear (Mendelson 2003; Coyne and Orr 2004; Mendelson et al. 2007; Sánchez-Guillén et al. 2012; Pulido-Santacruz et al. 2018). Preferences for conspecifics can evolve over short time scales and are influenced by social interactions through mechanisms such as sexual imprinting and learning (Verzijden et al. 2012), potentially limiting gene flow between divergent groups prior to the accumulation of additional reproductive barriers (Coyne and Orr 1989, 2004; Price and Bouvier 2002; Mendelson 2003). However, assortative mating can rapidly break down in response to changes in the ecological or social environment (Seehausen et al. 1997; Randler 2002; Secondi et al. 2014; Schumer et al. 2017) and may not be sufficient to prevent the collapse of divergent organisms (Irwin 2020). In contrast, post-mating barriers, such as genetic incompatibilities, usually take longer to accumulate than the typical time to speciation in many taxa, but are difficult to reverse once they evolve (Coyne and Orr 1989; Price and Bouvier 2002). Recent advances in genome sequencing technology have enabled the finer-resolution examination of the genomic architecture of divergence during incipient speciation and the specific genomic regions that may be involved in reproductive isolation between closely related organisms (Seehausen et al. 2014). However, few studies have combined fine-scale behavioral observations with emerging sequencing technology to examine the reproductive barriers that generate these genomic patterns early in the speciation process.

My dissertation sheds light on these gaps in our understanding of incipient speciation by identifying the traits that mediate genetic exchange between closely related species and

examining their roles in the promotion and/or maintenance of divergence in two recent avian radiations: the barn swallow (*Hirundo rustica*) species complex and the southern capuchino seedeater (*Sporophila*) radiation. Barn swallows comprise six subspecies that are distributed throughout the Palearctic and show substantial variation in migratory behavior and sexual signals (Scordato and Safran 2014). Two subspecies, *H. r. rustica* and *H. r. gutturalis*, share a hybrid zone in Gansu Province, China. Preliminary data from stable isotopes suggested that this contact zone likely represents a migratory divide, with *H. r. rustica* potentially migrating to the west and *H. r. gutturalis* to the east of the Tibetan Plateau en route to their respective wintering grounds. The southern capuchino seedeaters are finch-like Neotropical birds that, along with Darwin's finches, show exceptional diversification rates relative to other members of the tanager family (Burns et al. 2014). However, unlike Darwin's finches, capuchinos are thought to have diversified through sexual selection on melanin-based plumage traits rather than natural selection on foraging-related traits (Campagna et al. 2017; Lawson and Petren 2017). The capuchino radiation is characterized by high sympatry and extremely low levels of ecological and genomic divergence (mean genome-wide differentiation; $F_{ST} = 0.008$), despite striking differences in male plumage coloration and song used in mate attraction and territorial defense (Campagna et al. 2012, 2017). Seven of the ten capuchino members (*S. iberaensis*, *S. cinnamomea*, *S. palustris*, *S. hypochroma*, *S. hypoxantha*, *S. ruficollis*, and *S. pileata*) coexist in the grasslands of Iberá National Park in Corrientes, Argentina, providing the opportunity to directly examine patterns of genetic exchange between the phenotypically divergent species (Campagna et al. 2017).

Given their recent origin, both systems provide the opportunity to test the contribution of various behavioral and phenotypic traits to patterns of mate choice and genetic exchange early in the speciation process. In Chapter 2, I present a conceptual study that discusses how parallel

developments in animal tracking and genomic sequencing technology may be leveraged at migratory divides, or locations where co-occurring breeding populations pursue divergent strategies during the non-breeding season, to clarify the role of seasonal migration in the speciation process. Chapter 3 applies some of these ideas to a putative migratory divide between two barn swallow subspecies in Gansu Province, China, by combining high-throughput genomic sequencing with animal-borne tracking devices and stable isotopes to examine whether distinct migratory strategies correspond to differences in genome-wide ancestry between the subspecies. In Chapter 4, I use whole-genome sequencing and detailed behavioral observations and experiments to evaluate the role of pre-mating isolation in the maintenance of species boundaries between two species of capuchino seedeaters (*S. hypoxantha* and *S. iberaensis*), the genomic basis of the traits that limit interbreeding, and the origin of the genomic variants that differentiate the species in the context of the broader capuchino radiation. My results indicated that the reshuffling of standing genetic variation among closely related capuchino species likely facilitated rapid speciation and generated the remarkable phenotypic diversity observed in this group. One way in which genetic variation could be reshuffled is through meiotic recombination, a fundamental evolutionary process that can either facilitate divergence by generating novel genetic combinations or impede diversification by breaking apart favorable trait combinations in the presence of gene flow. Chapter 5 evaluates the contribution of variation in recombination rate to rapid speciation in capuchino seedeaters by examining the position of divergence peaks containing pigmentation genes relative to regions of high and low recombination. Finally, Chapter 6 synthesizes the results of my dissertation research and discusses the implications of this work.

CHAPTER 2

THE ROLE OF SEASONAL MIGRATION IN POPULATION DIVERGENCE AND REPRODUCTIVE ISOLATION[1]

2.1 ABSTRACT

Seasonal journeys between breeding and non-breeding habitat are undertaken by a diverse array of animals. Parallel developments in tracking and genomic methods are enabling finer resolution of these movements and their role in the evolutionary process. Evidence from allopatric and co-occurring breeding populations indicates that variation in migratory behavior is often associated with genetic differentiation. While assortative mating and selection against hybrids due to divergent migratory phenotypes can contribute to reproductive isolation, the details of these mechanisms remain unclear. Here we identify gaps in our understanding of the role of seasonal migration in the speciation process and propose a framework to test the relative significance of reproductive barriers associated with variation in migratory behavior that might underlie population differentiation.

2.2 EVOLUTIONARY OUTCOMES OF SEASONAL MIGRATION

Animal migration has evolved independently and repeatedly in a diverse array of taxonomic groups, including insects, fish, amphibians, birds, and mammals. Seasonal movements between breeding and non-breeding locations, hereafter 'seasonal migration,' provide numerous ecological benefits, allowing migratory organisms to escape competition,

[1] This chapter was published in *Trends in Ecology and Evolution* with Rebecca J. Safran and Elizabeth S. C. Scordato

colonize favorable habitats, and evade predators and parasites (Milner-Gulland et al. 2011). Although these seasonal journeys have captivated the human imagination for thousands of years (Aristotle *c*. 350 BC), until recently studies on migratory behavior heavily relied on mark–recapture efforts to gather information on the timing of migration and migratory routes of individuals (Webster et al. 2002). However, new innovations in tracking technology [reviewed in (Webster et al. 2002; Bridge et al. 2011)], particularly the ongoing miniaturization of tracking devices (*e.g.*, geolocators, satellite transmitters) and the development of indirect methods of inferring migratory behavior (*e.g.*, stable isotope analysis), have created unprecedented opportunities to examine variation in migratory strategies among a wide range of taxonomic groups (Kays et al. 2015). Combined with parallel advances in high-throughput genomic sequencing technology [see (Seehausen et al. 2014; Andrews et al. 2016)], these developments make possible previously intractable avenues of research on the evolutionary consequences of seasonal migration, particularly the contribution of seasonal movements to the maintenance or collapse of species boundaries (Shafer et al. 2016).

Seasonal migration requires an integrated suite of adaptations necessary for long-distance movement and survival in different environments (Milner-Gulland et al. 2011). As a result, migration can promote population differentiation by exposing lineages to divergent ecological conditions or limiting opportunities for genetic exchange through assortative mating or selection against hybridization (Winker 2010). Here we outline current knowledge about the role of seasonal migration in population divergence and provide a framework for studying how seasonal movements contribute to the evolution and maintenance of reproductive isolation. First, we discuss how divergent migratory behavior can lead to the evolution of differentiation between geographically isolated groups. Then, we describe how co-occurring breeding populations that

exhibit divergent migratory strategies present exciting opportunities to further our understanding of the contribution of seasonal migration to reproductive isolation.

2.3 LINKING MIGRATORY PHENOTYPE AND POPULATION DIFFERENTIATION IN ALLOPATRY

Research linking divergence in migratory phenotype to the speciation process has revealed that alterations in migratory behavior, combined with strong site fidelity, are often associated with genetic differentiation among closely related populations [*e.g.*, sharks (Jorgensen et al. 2010), sea turtles (Bowen et al. 1993), warblers (Irwin et al. 2011), salmon (Quinn et al. 2000)]. For example, Schreibers' long-fingered bat (*Miniopterus schreibersii natalensis*), which migrates between wintering roosts (hibernacula) and summer maternity colonies in South Africa, exhibits substantial genetic structure that corresponds to local habitats, wing morphology, and migratory behavior, suggesting that adaptation to local environmental conditions surrounding roosts could restrict gene flow between populations (Miller-Butterworth et al. 2003). In addition, the marked segregation of mtDNA haplotypes among subpopulations of humpback whales (*Megaptera novaeanglieae*) corresponds to patterns of seasonal migration, indicating a role of maternal fidelity to discrete migratory destinations in genetic differentiation (Baker et al. 1990).

While associations between migratory behavior and genetic divergence in allopatry indicate a possible role of seasonal migration in differentiation, it is difficult to disentangle the influence of migratory phenotype on the accumulation of genetic differentiation from the confounding effects of local adaptation and geographic distance. To address these challenges, we propose indirect methods of inferring the relative influence of migratory traits and geographic distance on patterns of genetic divergence (Table 2.1). These methods will allow researchers to

evaluate the potential role of divergent migratory phenotypes in population differentiation when reproductive barriers cannot be tested directly.

Table 2.1. Demonstrating that divergent migratory phenotypes influence genetic differentiation among allopatric populations (Steps 1-3) or pre-zygotic reproductive isolation at migratory divides (Steps 1-4). This framework should be applied to closely related groups in the early stages of the speciation continuum that exhibit alternative migratory strategies.

Step 1: Document differences in migratory phenotype between breeding populations. Alternative migratory phenotypes can be detected through one or more of the following methods: stable isotopes, tracking devices, orientation funnels, recapture records, and genetic markers. In a migratory divide, individuals that differ in migratory behavior should co-occur during the breeding season.
Step 2: Assess genetic divergence between populations that differ in migratory phenotype. The level of genetic differentiation can be quite low among recently isolated populations and at young migratory divides. We therefore advocate the use of high-throughput sequencing technology to generate the high-resolution genomic data necessary to detect regions of restricted gene flow extremely early in divergence. Association mapping for allopatric scenarios (Hecht et al. 2013) and admixture mapping in hybrid zones (Gompert et al. 2017) are both useful approaches for analyzing the genomic architecture of migratory behavior (Delmore and Liedvogel 2016). Several candidate genes (*e.g., CLOCK*, *ADCYAP1*) appear influential in the expression of alternative migratory phenotypes (Liedvogel et al. 2011; Delmore and Liedvogel 2016; Franchini et al. 2017). Determining whether candidate loci show elevated patterns of genomic differentiation relative to background levels of divergence between individuals with alternative migratory phenotypes can provide additional insight into the role of migratory behavior in the buildup of genetic divergence (Delmore et al. 2015a).
Step 3: Assess trait correlations (*e.g.*, timing of migration, migratory route, phenotypic differences, and habitat selection). As geographic isolation can yield divergence along numerous phenotypic axes, and many migratory divides have formed through secondary contact, it is important to identify and distinguish between traits that truly underlie patterns of mate choice and traits that are associated with the actual target of mate selection and therefore appear to influence patterns of genetic exchange. For allopatric scenarios, we advocate the use of variance partitioning to infer the role of migratory phenotype in genetic differentiation by examining the relative contributions of migration-related traits, traits unrelated to migratory behavior, and geographic distance to pairwise genetic divergence (Safran et al. 2016). Multiple matrix regression analyses can be used to partition variation among correlated variables (*e.g.,* spatial autocorrelation (Wang 2013)).
Step 4: Evaluate the mechanistic basis of assortative mating and demonstrate its link to migratory behavior (only possible at migratory divides). We offer examples of correlational tests and experimental manipulations that can be used to tease apart the roles of multiple associated traits in assortative mating (Table 2.2). However, we recognize that experimental tests will only be feasible in a small number of study systems.

2.4 LINKING MIGRATORY PHENOTYPE AND POPULATION DIFFERENTIATION AT MIGRATORY DIVIDES

Divergent migratory phenotypes are often associated with genetic differentiation between closely related allopatric populations (Bowen et al. 1993; Quinn et al. 2000; Jorgensen et al. 2010; Irwin et al. 2011); however, we lack information about the direct contribution of divergent migratory phenotypes to reproductive isolation. Migratory divides provide a unique opportunity to address this gap by enabling researchers to directly test the strength of reproductive barriers linked to seasonal migration. While the term migratory divide lacks a consistent definition in the literature and has largely been restricted to avian systems, we use the phrase broadly to refer to populations that coexist during the breeding season but overwinter in separate locations [*e.g.*, anadromous vs non-anadromous forms of salmonid fish (Dodson et al. 2013), sedentary and migratory herds of caribou (McDevitt et al. 2009), wolves that prey on caribou with divergent migratory strategies (Musiani et al. 2007)] or follow different migratory routes to the non-breeding grounds (*e.g.*, beluga whales that migrate to distinct summering areas following reproduction (O'Corry-Crowe et al. 1997), storks that take different migratory routes to bypass the Mediterranean Sea (Bobek et al. 2008)). Migratory divides are formed when divergent migratory strategies that arise in allopatry are maintained on secondary contact (Figure 2.1A) or when divergent migratory behavior evolves in situ through primary divergence (Figure 2.1B).

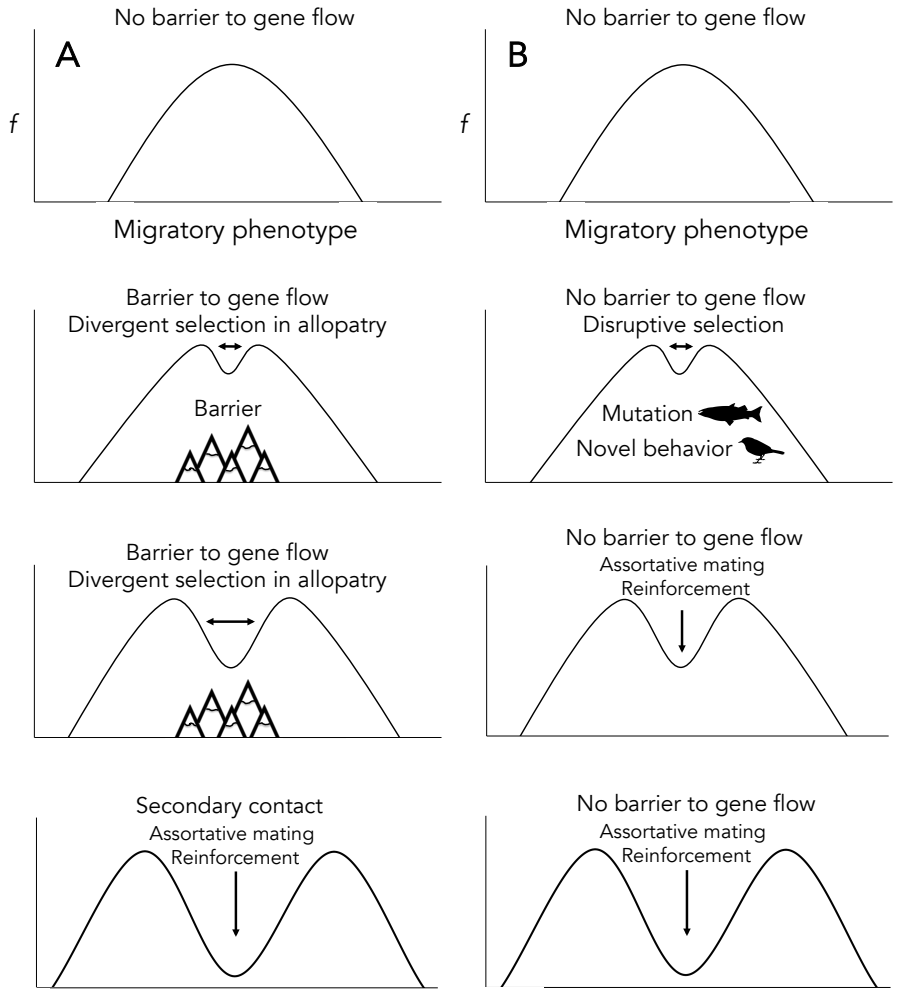


Figure 2.1. A diagram outlining how **(A)** secondary contact and **(B)** primary divergence could lead to reproductive isolation by migratory phenotype and the eventual buildup of genetic differentiation at migratory divides. Top panels represent panmictic populations with standing genetic variation in migratory phenotype. (**A**) Exposure to divergent ecological conditions and physical barriers to migration can lead to the evolution of divergent migratory strategies in geographically isolated populations. Depending on the duration of isolation in allopatry, alternative migratory forms may or may not accumulate significant genetic differences prior to secondary contact. When populations are brought into sympatry, existing trait differences might be sufficient to yield assortative mating by migratory phenotype. Otherwise, selection against hybrids could lead to reinforcement, further promoting trait divergence and the buildup of genetic differentiation over time between individuals with alternative migratory phenotypes. (**B)** In primary divergence, a novel mutation affecting migration (*e.g.,* salmon (Prince et al. 2017)) or a new migratory behavior (*e.g.,* blackcaps (Berthold et al. 1992)) can arise in a previously homogenous population. If individuals with different migratory phenotypes have equally high reproductive fitness, divergent migratory strategies will be maintained in sympatry. However, assortative mating and/or selection against hybrids, if present, will cause alternative migratory phenotypes to diverge from one another over time in traits involved in migration and eventually lead to the accumulation of genetic divergence between migratory forms.

In birds, migratory divides are often associated with significant genetic differentiation (Rolshausen et al. 2009; Delmore et al. 2016), formed through secondary contact (Ruegg 2008; Bensch et al. 2009), and hypothesized to directly promote and/or maintain reproductive isolation through prezygotic and post-zygotic isolating mechanisms (Irwin and Irwin 2005). In the case of prezygotic reproductive isolation, differential timing of arrival on the breeding grounds between individuals with divergent migratory phenotypes can lead to assortative mating (temporal isolation) (Bearhop et al. 2005). Furthermore, selection against hybrids that undertake inferior migratory routes, inherit maladaptive trait combinations, or experience suboptimal life history schedules associated with their migratory phenotype might limit gene flow through post-zygotic reproductive isolation (Rohwer and Irwin 2011; Delmore and Irwin 2014). For example, hybrids between populations that follow divergent migratory routes around a geographic barrier could experience increased mortality if they travel directly across the barrier during migration. Finally, post-zygotic selection against hybrids could lead to stronger prezygotic barriers via reinforcement (Irwin and Irwin 2005). In this scenario, selection for increased mate discrimination to avoid maladaptive hybridization between distinct migratory forms would yield greater divergence in traits associated with mate choice. Nonetheless, a comparative analysis of North American birds found that sister pairs that differ in migratory behavior exhibit lower rates of phenotypic divergence, suggesting that seasonal movements might be sufficient to maintain reproductive isolation even in the absence of phenotypic differentiation (Delmore et al. 2015b).

Divergent migratory phenotypes can also indirectly influence reproductive isolation through carry-over effects from the non-breeding season or differences in selection pressures between populations (Saino et al. 2004). For example, exposure to divergent ecological conditions on the non-breeding grounds, such as the availability of dietary resources, the sensory

environment in which signaling occurs, and the composition of predator and parasite communities, as a result of alternative migratory phenotypes could indirectly lead to divergence in traits involved in assortative mating, as sexual signals can rapidly diverge in response to ecological selection (Maan and Seehausen 2011). In addition, conflicting selection processes that arise as a byproduct of variation in migratory phenotype could promote divergence in signal traits and sexual preferences between populations (Irwin and Irwin 2005). For instance, comparative analyses of sedentary and migratory populations of several avian species have found that migratory males produce longer, female-directed songs while resident males produce shorter, repetitive songs effective in territorial interactions, which could potentially yield assortative mating on secondary contact (Irwin 2000; Collins et al. 2009).

A disproportionate amount of research on migratory divides has been conducted in avian systems, which are often characterized by strong breeding site fidelity, non-random mating, and variation in seasonal migration. Other taxonomic groups undergo migratory journeys that span multiple generations [*e.g.*, insects (Milner-Gulland et al. 2011)], choose mates, reproduce, and raise offspring in geographically distinct locations [*e.g.*, bats (Miller-Butterworth et al. 2003)] and depend to various degrees on cultural transmission to undertake successful migratory journeys [*e.g.*, wildebeest (Milner-Gulland et al. 2011)]. We currently lack sufficient information on migratory behavior to predict how divergent migratory phenotypes may influence patterns of genetic exchange in many systems with diverse life history strategies. However, the joint application of emerging genomic and tracking technologies in understudied systems will shed light on the prevalence of migratory divides and their broad importance to reproductive isolation and population divergence across animals as a whole.

2.5 MIGRATORY DIVIDES: REMAINING QUESTIONS

While a variety of mechanisms are thought to underlie reductions in gene flow at migratory divides, the relative significance of possible isolating mechanisms, including pre- and post-zygotic barriers, in the evolution of reproductive isolation remains unclear. Recent advances in tracking techniques and high-throughput sequencing technology are making it possible to test these putative mechanisms for the first time by allowing researchers to measure variation in migratory behavior and link these movement patterns to gene flow and genomic ancestry in contact zones (Kays et al. 2015; Shafer et al. 2016).

To advance our understanding of the evolutionary consequences of seasonal migration as new migratory divides continue to be discovered (Bobek et al. 2008; Davis et al. 2016), we: (i) highlight remaining questions regarding the link between migratory phenotype and reproductive isolation by examining five case studies; (ii) propose a conceptual framework for studying the influence of divergent migratory phenotypes in the evolution of prezygotic isolation; and (iii) discuss how parallel developments in animal tracking and genomic sequencing technology can be leveraged to track patterns of gene flow as a function of migratory behavior and infer post-zygotic selection against hybrids at migratory divides.

Migratory divide case studies

In Table 2.2, we summarize five case studies drawn from a variety of taxonomic groups to point out gaps in our understanding of how divergent migratory phenotypes contribute to reproductive isolation at migratory divides. In particular, few studies have investigated patterns of gene flow between individuals with divergent migratory phenotypes or matched observations of pairing data to paternity. Analyzing both within-pair and extra-pair mating could be revealing;

for example, if rates of extra-pair paternity are high, and individuals mate assortatively by migratory phenotype with social but not extra-pair mates, studies that fail to track patterns of extra-pair mate choice could overestimate the importance of migratory behavior for reproductive isolation. In the following section, we present a framework to address these gaps by explicitly analyzing whether divergent migratory phenotypes lead to prezygotic isolation at migratory divides. While populations with divergent migratory strategies often differ in features that affect mating decisions, such as timing of breeding (Bearhop et al. 2005), phenotype (Ruegg 2008), and habitat choice (Rolshausen et al. 2013), comparatively little is known about the prevalence of maladaptive genotypes in hybrids, making assessment of post-zygotic barriers challenging.

Table 2.2. **Migratory divide case studies.** We examine five case studies of migratory divides to highlight the persisting gaps in our understanding of how divergent migratory phenotypes contribute to the generation and maintenance of pre-zygotic and post-zygotic reproductive isolation (RI).

System	Description	Biogeographic history	Axes of divergence	Maintenance of RI	Gaps
Eurasian blackcap (*Sylvia atricapilla*)	Southwest (SW) migrants travel to the Iberian Peninsula & Africa; northwest (NW) migrants travel to the United Kingdom (UK) & Iceland from sympatric breeding grounds in Germany & Austria (Berthold et al. 1992)	Primary divergence: novel NW direction evolved in the 1960s due to warmer winter conditions & supplemental feeding in the UK (Berthold et al. 1992; Rolshausen et al. 2010)	Genetic & phenotypic divergence (wing shape, beak width, & plumage coloration) (Rolshausen et al. 2009); timing of arrival on breeding grounds (Bearhop et al. 2005)	Pre-zygotic: assortative mating by timing of arrival (Bearhop et al. 2005), but SW-migrants outnumber NW-migrants on the breeding grounds, & females arrive much later than males (Rolshausen et al. 2010); assortative mating by habitat selection on the breeding grounds (Rolshausen et al. 2013) Post-zygotic: exhibit continuous wintering distribution from SW to NW Europe (Irwin 2009), but hybrids may orient in intermediate direction (Helbig 1996)	No study has examined patterns of extra-pair mate choice or selection against hybrids
Sockeye salmon (*Oncorhynchus nerka*)	Anadromous 'sockeye' migrate to the Pacific Ocean to mature; non-anadromous 'kokanee' remain in lakes until maturity; both ecotypes often spawn in the same freshwater bodies (Wood et al. 2008)	Primary divergence: non-anadromous kokanee repeatedly evolved from anadromous sockeye (Wood et al. 2008)	Genetic & phenotypic divergence (body size); selection of oviposition sites (Dodson et al. 2013); ability to utilize carotenoids (Craig and Foote 2001)	Pre-zygotic: assortative mating by body size (Foote and Larkin 1988); habitat isolation due to disruptive selection on female oviposition sites (Dodson et al. 2013); sexual selection against green-colored hybrids (Dodson et al. 2013)) Post-zygotic: sockeye and kokanee produce viable, fertile offspring in hatcheries (Craig and Foote 2001); potential pleiotropic effects of carotenoid sequestration on smoltification in kokanee (Dodson et al. 2013)	Few studies have evaluated the mechanisms maintaining reproductive isolation between distinct migratory forms; pre-zygotic mechanisms appear important but patterns of genetic exchange between ecotypes are unclear

Willow warbler (*Phylloscopus trochilus*)	*P. t. acredula* migrates southeast to southern Africa & *P. t. trochilus* migrates southwest to west Africa from sympatric breeding grounds in Scandinavia (Chamberlain et al. 2000).	Secondary contact following postglacial expansion from a common refuge population (Bensch et al. 2009)	Little genetic differentiation at neutral loci (Bensch et al. 2009) but divergence at 2 AFLP-derived markers; phenotypic divergence (body size, plumage coloration); timing of arrival (Bensch et al. 1999)	Pre-zygotic: assortative mating proposed by timing of arrival (Bensch et al. 1999); no assortative mating between social mates by coloration, morphology, or isotopes (Liedvogel et al. 2014); hybridization is widespread (Liedvogel et al. 2014) Post-zygotic: selection proposed against hybrids that take maladaptive migratory routes (Liedvogel et al. 2014)	Post-zygotic selection against hybrids proposed as the only force that could explain the migratory divide but no study has examined extra-pair mate choice, hybrid fitness, or the migratory routes of hybrid individuals
Atlantic eels (*Anguilla rostrata* & *Anguilla anguilla*)	Larval *A. rostrata* & *A. anguilla* migrate on ocean currents from overlapping spawning grounds to streams & lakes in North America & Europe, respectively (Albert et al. 2006)	Secondary contact (Albert et al. 2006)	Genetic (Jacobsen et al. 2014) & morphological divergence (vertebrae count) (Albert et al. 2006); larval stage duration (Jacobsen et al. 2014); spawning time (Pujolar et al. 2014)	Pre-zygotic: assortative mating proposed through behavioral or ecological mechanisms (*e.g.*, timing of spawning (Pujolar et al. 2014)) Post-zygotic: hybrids make up a higher proportion of later life history stages (potentially due to increased hybrid survival) (Albert et al. 2006), but could experience lower fitness outside of hybrid zone; possibility of genetic incompatibilities	No studies have tracked patterns of genetic exchange on spawning grounds or hybrid fitness
Swainson's thrush (*Catharus ustulatus*)	*C. u. ustulatus* migrates along the coast to southern Mexico and Central America, *C. u. swainsoni* migrates inland to overwinter from Panama to Argentina (Ruegg and Smith 2002)	Secondary contact following postglacial expansion from separate glacial refugia (Ruegg 2008)	Genetic & phenotypic divergence (body size, plumage coloration, song); breeding habitat (Ruegg et al. 2006; Ruegg 2008); timing of arrival (Ruegg et al. 2012); migratory orientation strongly associated with a genomic region of elevated differentiation (Delmore et al. 2016)	Pre-zygotic: assortative mating proposed by timing of arrival, song, or habitat selection on breeding grounds (Ruegg et al. 2012) Post-zygotic: hybrids take intermediate migratory routes, but return rates are similar between hybrids and parentals (Delmore and Irwin 2014)	No study has examined social or extra-pair mate choice or hybrid fitness at the divide; effect of intermediate migratory strategies for hybrid survival is unclear

2.6 A CONCEPTUAL FRAMEWORK FOR STUDYING PRE-ZYGOTIC ISOLATION AT MIGRATORY DIVIDES

Geographic context and the promotion and/or maintenance of reproductive isolation

Many migratory divides in birds and mammals have formed through secondary contact between populations that evolved divergent migratory phenotypes in allopatry (Ruegg 2008; Bensch et al. 2009; McDevitt et al. 2009); however, several divides, particularly in fish, are likely to have evolved through primary divergence (Table 2.2). We outline how secondary contact (Figure 2.1A) and primary divergence (Figure 2.1B) can lead to pre- and/or post-zygotic isolation by migratory phenotype and the eventual accumulation of genetic differences between populations. Regardless of the geographic context of divergence, the long-term expectation is largely the same if divergent migratory phenotypes contribute to prezygotic reproductive isolation at migratory divides. In divides formed through either primary divergence or secondary contact, traits involved in assortative mating by migratory phenotype will eventually exhibit a non-unimodal (*e.g.*, bimodal) distribution, with little overlap between migratory forms, and individuals that pursue alternative migratory strategies will become genetically differentiated from one another. The rate at which genetic divergence accumulates will depend on the history of geographic isolation, the strength of assortative mating and selection against hybrids, the genetic architecture of migratory behavior, and the roles of social learning and phenotypic plasticity in generating variation in migratory phenotype. For recently evolved migratory divides, other measures, such as mate preferences, pairing decisions, and patterns of paternity, can be used to detect reproductive isolation in lieu of genetic divergence.

The application of emerging technology to collect fine-scale data on variation in migratory routes, geographic location and habitat during the non-breeding season, and timing of

arrival on the breeding grounds will facilitate the documentation of traits involved in prezygotic isolation. While many studies have used low-resolution genomic markers, such as micro-satellites and allozymes, to detect regions of restricted gene flow between organisms with alternative migratory phenotypes (Lyons et al. 2012; Dodson et al. 2013), these markers lack the resolution necessary to detect genetic differentiation extremely early in divergence. We advocate the use of high-throughput sequencing technology to produce the high-resolution data necessary to examine the accumulation of genetic differentiation between closely related populations and identify the genomic regions underlying divergent migratory phenotypes (Delmore et al. 2016).

The influence of cultural transmission and phenotypic plasticity on the accumulation of genetic differentiation at migratory divides is largely unknown. While both environmental and genetic drivers of alternative migratory phenotypes should eventually lead to genetic differentiation if individuals mate assortatively by migratory pattern, the effect of these processes on the rate at which genetic divergence accumulates is poorly understood. Cross-fostering experiments between individuals that differ in migratory behavior and the examination of correlations between genomic ancestry and migratory phenotype will allow researchers to investigate how mode of inheritance influences the strength of reproductive isolation and infer the relative contributions of genetics, learning, and phenotypic plasticity to the generation of alternative migratory strategies.

Testing the mechanistic basis of assortative mating following secondary contact

Distinct migratory forms have often diverged in numerous traits in allopatry that could mediate assortative mating on secondary contact. Associations between traits due to divergence in allopatry can yield misleading conclusions about the contribution of seasonal migration to

prezygotic isolation. To evaluate the role of migratory phenotype *per se* in the maintenance of reproductive isolation, we emphasize the need to compare allopatric and sympatric populations of the same species and comprehensively assess associations between diverse traits, such as timing of migration, propensity to migrate, morphological and phenotypic traits, habitat selection, and genomic ancestry (Table 2.1).

Once associations between behavioral and phenotypic traits have been identified, studies should attempt to tease apart these associations through statistical analyses and experimental tests. Observing changes in patterns of hybridization after an experimental manipulation is a particularly powerful way to detect a causal association between the manipulated trait and barriers to reproduction. This step will allow researchers to evaluate the mechanistic basis of assortative mating and its link to migratory behavior. In particular, sexually selected and naturally selected phenotypic traits (Irwin 2000; Collins et al. 2009), timing of arrival (Bearhop et al. 2005), habitat selection on the breeding grounds (Rolshausen et al. 2013), and genomic ancestry could underlie assortative mating at migratory divides. In Table 2.3, we propose a series of correlational tests and experimental manipulations that will allow researchers to: (i) identify potentially important prezygotic isolating mechanisms that could maintain divergent migratory phenotypes at divides; and (ii) break apart trait correlations to infer the relative importance of these possible mechanisms of assortative mating. While certain systems are more amenable to experimental manipulation than others, correlational analyses alone will provide substantial insight into the traits underlying assortative mating at migratory divides. The application of this framework will shed light on the frequency with which migratory behavior directly (*e.g.*, through assortative mating by timing of arrival) and indirectly (*e.g.*, through carry-over effects from the non-breeding season) promotes reproductive isolation across a diversity of taxonomic groups.

Table 2.3. Evaluating the mechanistic basis of assortative mating at migratory divides. Here, we identify potential non-mutually exclusive mechanisms of assortative mating (AM) at migratory divides and consider how divergent migratory phenotypes can lead to pre-zygotic reproductive isolation within sympatric breeding populations. We propose correlational tests and, when feasible, follow-up experimental manipulations to infer the relative importance of possible mechanisms of assortative mating. The joint application of emerging tracking and genomic sequencing technologies will allow researchers to accurately assess and link various aspects of seasonal migration, including timing of arrival on the breeding grounds, non-breeding location (and potential carry-over effects from the non-breeding season), and genomic ancestry, to pairing decisions and patterns of paternity on the breeding grounds.

Basis of assortative mating [AM]	**Prediction**	**Correlational test**	**Experimental manipulation**
Phenotype - sexual selection	Mate selection maintains AM by migratory behavior (*e.g.*, signal traits predict migratory behavior)	Association between signal traits and migratory behavior, correlation between signal traits of mates	Remove or enhance signal differences and analyze patterns of genetic exchange (*e.g.*, conduct mate choice trials with manipulated pheromone blends (Moore et al. 2001))
Phenotype- natural selection	Phenotype maintains AM by migratory behavior (*e.g.*, wing length predicts migratory behavior)	Association between phenotypic traits and migratory behavior, correlation between phenotypic traits of mates	Remove or enhance trait differences and analyze patterns of genetic exchange (*e.g.*, alter pigmentation patterns involved in background matching (Karpestam et al. 2012))
Behavior - timing	Timing differences maintain AM by migratory behavior	Association between timing of arrival and migratory behavior, correlation between migratory timing of mates	Remove timing differences and analyze patterns of genetic exchange (*e.g.*, delay breeding of early-arriving males (McKellar et al. 2013))
Behavior - habitat	Habitat selection maintains AM by migratory behavior	Correlation between wintering habitat and habitat selection on breeding grounds	Remove or enhance habitat differences and analyze patterns of genetic exchange (*e.g.*, alter the stone cover on male territories (Taborsky et al. 2014))
Ancestry	Selection of genetically similar mates maintains AM by migratory behavior	Higher genetic relatedness between mates than expected by chance	Force heterotypic matings to expose post-zygotic isolating mechanisms (*e.g.*, conduct hybrid crosses (Williams and Mendelson 2014))

2.7 DETECTING POST-ZYGOTIC ISOLATION DUE TO DIVERGENT MIGRATORY PHENOTYPES

Post-zygotic isolation can also restrict gene flow between groups that pursue alternative migratory strategies and might be particularly prevalent in certain systems, such as organisms that reproduce through broadcast spawning. To date, post-zygotic isolation as a function of migratory phenotype has been considered almost exclusively in songbirds, which, due to their small size, cannot support tracking devices that continuously transmit data in real time. As a result, researchers have attempted to infer post-zygotic barriers to reproduction by tracking the migratory routes of hybrid individuals using orientation funnels and geolocators, which must be retrieved to obtain tracking data (Helbig 1996; Delmore and Irwin 2014). In addition, studies have analyzed the return rates of hybrids relative to parentals (Delmore and Irwin 2014) and compared the width of genetic and phenotypic clines associated with migratory divides to neutral expectations (Ruegg 2008; Bensch et al. 2009). However, it is impossible to pinpoint the source of selection against hybrids and the role of seasonal migration *per se* in post-zygotic isolation using these approaches. In addition, documenting that hybrids follow intermediate migratory routes relative to parentals does not necessarily imply lower hybrid fitness. Finally, cultural inheritance of migratory route could limit the strength of post-zygotic barriers to reproduction in certain systems (Milner-Gulland et al. 2011).

To investigate sources of post-zygotic isolation in more detail, we encourage research in non-avian systems, such as fish, caribou, and other large animals that can be tracked using acoustic telemetry or support GPS-based devices that transmit real-time data on animal positions. The high-resolution tracking data produced by these technologies will provide information on mortality rates, the avoidance of geographic barriers (*e.g.*, deserts, mountain ranges) during

terrestrial migrations, and the intermediate nature of migratory routes in hybrids, improving our ability to draw inferences about the strength of post-zygotic isolation due to divergent migratory phenotypes. Combined with detailed investigations into patterns of genetic exchange and hybrid fitness on the breeding grounds, fine-scale spatial data collected during the non-breeding season will provide insight into factors other than migratory route (*e.g.*, carry-over effects from the non-breeding grounds) that promote selection against hybrids. In certain systems it may also be possible to conduct experimental crosses between individuals with alternative migratory phenotypes to examine inviability, sterility, and other sources of post-zygotic selection. Eventually, the further miniaturization of tracking devices, particularly those with GPS technology, and the development of satellite-based infrastructure [*e.g.*, the ICARUS (International Cooperation for Animal Research Using Space) Initiative] to track the movements of small animals will open new opportunities to test the prevalence of post-zygotic isolation as a function of migratory phenotype in a wide range of taxonomic groups.

2.8 CONCLUDING REMARKS AND FUTURE DIRECTIONS

Migratory divides represent a unique opportunity to improve our understanding of the role of seasonal migration in speciation; however, we lack a conceptual framework for demonstrating the importance of divergent migratory phenotypes in pre-zygotic isolation. Here, we provide an overview of the numerous ways in which seasonal migration can contribute to genetic divergence among allopatric populations and reproductive isolation in areas of geographic overlap. We advance a conceptual framework to evaluate the relative significance of pre-zygotic isolating mechanisms that can underlie population differentiation due to variation in migratory behavior. In addition, we describe methods of inferring the contribution of post-

zygotic barriers to reproductive isolation between individuals with divergent migratory phenotypes. Our framework discusses opportunities to leverage emerging technology in a wide range of taxonomic systems to track patterns of genetic exchange between individuals that pursue divergent migratory strategies and offers guidelines to explicitly examine the link between seasonal migration and the accumulation of reproductive isolating barriers. The generation of high-resolution spatial and genomic data advocated in this approach will not only provide direct insight into the role of migratory phenotype in the evolution and maintenance of reproductive isolation, but also inform other important aspects of seasonal migration, such as the repeatability of individual migratory routes, the genetic basis of migratory behavior (Liedvogel et al. 2011; Franchini et al. 2017), and the roles of phenotypic plasticity, social imprinting, and genetic variation in the generation of alternative migratory phenotypes. Growing evidence suggests that climate change can impact species distributions, influence the timing of migration, and select for the evolution of residence in certain migratory populations (Pulido and Berthold 2010). Understanding the lability of migratory behavior and the conditions under which seasonal migration promotes, maintains, or erodes barriers to genetic exchange in diverse taxonomic groups will become increasingly important as rapidly changing environmental conditions alter the journeys of migratory organisms in unpredictable ways.

CHAPTER 3

A MIGRATORY DIVIDE SPANNING TWO CONTINENTS IS ASSOCIATED WITH GENOMIC AND ECOLOGICAL DIVERGENCE[2]

3.1 ABSTRACT

Divergent migratory behavior can promote population differentiation by exposing organisms to different environments and restricting gene flow between taxa. Migratory divides provide an opportunity to evaluate the role of seasonal migration in population divergence, particularly the association between population structure and distinct migratory strategies. However, few studies have leveraged emerging technology at migratory divides to examine this association in detail. We combine light-level geolocators, genomic sequencing, and stable isotopes to document a striking migratory divide that coincides with genomic differentiation across a hybrid zone between barn swallow (*Hirundo rustica*) subspecies in China. Individuals on the western side, with predominately *H. r. rustica* ancestry, had comparatively enriched feather isotope values and overwintered in eastern Africa, while birds breeding on the eastern side, with *H. r. gutturalis* ancestry, had depleted isotope values and migrated to southern India. The two subspecies migrated using divergent routes around the high-altitude Karakoram Range and arrived on their breeding grounds over three weeks apart. These results indicate that assortative mating by timing of arrival and/or selection against hybrids that inherit intermediate migratory traits may limit interbreeding between the subspecies, and that inhospitable geographic

[2] This chapter was submitted to *Evolution* with Drew R. Schield, Elizabeth S.C. Scordato, Andrea Contina, Xin-Wei Da, Yang Liu, Yu Liu, Emilio Pagani-Núñez, Qing-Miao Ren, Chris C. R. Smith, Craig A. Stricker, Michael Wunder, David M. Zonana, and Rebecca J. Safran

features may have contributed to the diversification of Asian avifauna by influencing migratory patterns.

3.2 INTRODUCTION

Every year, billions of fish, mammals, insects, and birds undergo spectacular journeys in pursuit of seasonal changes in resource abundance that carry them vast distances from their breeding grounds (Dingle 2014). These annual migrations between breeding and nonbreeding habitats require an integrated suite of morphological, physiological, sensory, and behavioral adaptations necessary for long-distance movement, as well as survival and reproduction in disparate environments (Milner-Gulland et al. 2011; Dingle 2014). Given the complex nature of seasonal migration, divergence in migratory strategies between closely related taxa has been proposed to promote genetic divergence and population differentiation (Irwin and Irwin 2005; Winker 2010; Turbek et al. 2018). However, the difficulty of tracking seasonal movements across space and time has historically limited our understanding of the ways in which seasonal migration can contribute to patterns of population divergence (Kays et al. 2015).

Migratory divides—contact zones between breeding populations with divergent migratory strategies during the non-breeding season (Bensch et al. 1999; Delmore et al. 2012; Alvarado et al. 2014)—offer a unique opportunity to study the role of migratory behavior in reproductive isolation and speciation (Turbek et al. 2018). An overarching prediction regarding the contribution of divergent migratory strategies to population differentiation at migratory divides is an association between population structure, patterns of gene flow, and distinct migratory strategies. In these regions, divergent migratory behavior is thought to restrict genetic exchange through assortative mating by timing of arrival on the breeding grounds and/or selection against hybrids that inherit intermediate migratory traits (Bearhop et al. 2005; Irwin and

Irwin 2005; Delmore and Irwin 2014), leading to an accumulation of genomic divergence over time (Rolshausen et al. 2009). However, few studies have leveraged emerging animal tracking and genomic sequencing technology in migratory divides to study the relationship between genomic differentiation and divergent migratory strategies.

In Asia, the high-elevation Qinghai-Tibetan Plateau has been proposed as a hostile geographic barrier to animal migration (Delany et al. 2017; Liu et al. 2018) that has potentially contributed to the divergence of Asian avifauna (Irwin and Irwin 2005). According to compiled data from banding records and species range distributions, a disproportionate number of passerine species that breed in Siberia migrate either to the east or west of the Qinghai-Tibetan Plateau en route to their wintering grounds in southern Asia (Irwin and Irwin 2005). In addition, many species that circumvent the barrier in both directions exhibit migratory divides between eastern and western forms potentially associated with reproductive isolation on the breeding grounds (Irwin and Irwin 2005). By influencing the migratory routes of passerine species, major geographic features, such as the Qinghai-Tibetan Plateau and the nearby low-altitude Gobi and Taklamakan deserts, may have shaped the biogeographic patterns of songbirds in northern Asia. Nonetheless, no study has used direct animal-borne tracking methods to examine how migratory passerines navigate the Qinghai-Tibetan Plateau and similarly inhospitable features in central Asia (Delany et al. 2017).

Barn swallows (*Hirundo rustica*) are widely distributed throughout the Palearctic and comprise six subspecies that exhibit substantial variation in migratory behavior (Turner 2010). Two subspecies, *H. r. rustica*, which breeds throughout northern Africa, Europe, and western Asia, and *H. r. gutturalis*, which breeds in southern and eastern Asia, share a hybrid zone in northwestern China that was likely formed through secondary contact (Scordato et al. 2020).

Previous work, which used genotyping-by-sequencing (GBS) to examine patterns of divergence in barn swallows across Asia, documented low genomic differentiation (mean genome-wide F_{ST} = 0.028) between the subspecies and detected later generation hybrids and backcrossed individuals but few F1s in the hybrid zone in Gansu Province, China, providing evidence for limited but ongoing gene flow between *H. r. rustica* and *H. r. gutturalis* (Scordato et al. 2020). The study also found little evidence of within-population pre-mating isolation between the subspecies and indicated that parental individuals of *H. r. rustica* and *H. r. gutturalis* are largely confined to either end of the hybrid zone and do not overlap in the zone's center, which is composed of admixed individuals (Scordato et al. 2020). Thus, extrinsic selection against hybrids may contribute to the maintenance of subspecies boundaries between *H. r. rustica* and *H. r. gutturalis*. Finally, the subspecies exhibit substantial divergence in the stable-carbon isotope values of their tail feathers, which are molted on the wintering grounds (Turner 2010; Scordato et al. 2020). Stable isotopes values [*e.g.,* carbon ($\delta^{13}C$)], are incorporated into metabolically inert tissues, such as feathers, through diet and remain constant across the annual cycle, reflecting the isotopic composition of the local environment in which the feathers were grown (Hobson 1999; Rubenstein and Hobson 2004). Carbon isotopes, in particular, are influenced by the relative abundance of C3 and C4 plants in the area of feather growth (Hobson 1999; Rubenstein and Hobson 2004). Thus, strong divergence in $\delta^{13}C$ values between the subspecies likely indicates the presence of a migratory divide (Scordato et al. 2020).

To investigate the possibility that a migratory divide between *H. r. rustica* and *H. r. gutturalis* may be associated with subspecies boundaries, we generated higher-resolution whole-genome sequencing (WGS) data and attached light-level geolocators to barn swallows breeding on either end of the hybrid zone. These animal-borne tracking devices recorded light intensity

levels at specific intervals throughout the annual cycle. In addition, we used double digest restriction-site associated DNA (ddRAD) sequencing and stable-carbon ($\delta^{13}C$) and hydrogen ($\delta^{2}H$) isotope analysis to examine a larger sample of individuals concentrated in the area of putative hybridization and verify patterns observed in the geolocator data. Using these datasets, we analyzed whether distinct migratory strategies (1) corresponded to differences in genome-wide ancestry between the subspecies, as expected if migratory behavior contributes to reproductive isolation on the breeding grounds, and (2) coincided with geographic features proposed as major barriers to migration in Asian songbirds.

3.3 MATERIALS AND METHODS

Field sampling

We sampled 231 individuals of six barn swallow subspecies across eight sampling trips that took place during the barn swallow breeding season (April-July 2013 in Russia, April-July 2014 in China, Mongolia, and Japan, January-June 2015 in Israel, April 2015 in Egypt, May-June 2015 in China and Colorado, April-May 2016 in Morocco, July 2016 in China, and May 2017 in China). In each region, we captured barn swallows with mist nets, fit each bird with an individually numbered aluminum band, collected the two inner-most tail rectrices, and took a blood sample from the brachial vein prior to release. Blood samples were stored in lysis buffer and DNA was extracted with the DNeasy blood and tissue kit (Qiagen, CA, USA) and included in subsequent genomic analyses to improve the accuracy of variant calling within the hybrid zone.

Genome-wide ancestry: Whole-genome sequencing

We performed whole-genome sequencing on 168 individuals from allopatric populations of the six subspecies, as well as areas of putative hybridization, following the protocol detailed in Smith et al. (2018). Details on sample collection can be found in Table S3.1. Libraries were prepared at the University of Colorado Boulder BioFrontiers Institute using the ILLUMINA NEXTERA XT kit with version 2 dual indexes at 0.5× volume. Individual libraries were then pooled and size-selected for 300-700 bp fragments using a Pippin Prep. We sequenced the pooled libraries on two replicate lanes of an Illumina NovaSeq at the Novogene Genome Sequencing Company (Chula Vista, CA, USA) to generate 150-bp paired-end reads.

We used Trimmomatic (version 0.36) to remove bases at the ends of each read with quality scores below 30 and trim reads with an average quality score less than 30 or length less than 50 bp (Bolger et al. 2014). We then aligned the trimmed reads to a barn swallow reference genome (Formenti et al. 2018) with BWA-MEM (version 0.7.12) and called variants with SAMtools (version 1.5) and BCFtools (version 1.5) (Heng et al. 2009; Li 2013). To avoid sex-linked loci, we filtered out sex chromosomes by aligning the barn swallow scaffolds to flycatcher autosomes (Ellegren et al. 2012) and the chicken W chromosome (International Chicken Genome Consortium 2004) using the program MashMap (Jain et al. 2017, 2018). We set the mapping segment length parameter to 10,000 for scaffolds shorter than 1 Mb and 50,000 for scaffolds longer than 1 Mb. Like other alignment programs, MashMap outputs many short alignment segments. We assigned individual scaffolds to a chromosome if (1) more than half of the alignment segments landed on the same chromosome, and (2) more than half of the total scaffold length aligned to the same chromosome, or else excluded the scaffold from further analysis. For the retained barn swallow scaffolds, we approximated the base pair position by

subtracting half of the scaffold length from the median alignment position among the short alignments mapping to the assigned chromosome.

We restricted the final dataset to *H. r. rustica* and *H. r. gutturalis* individuals that were sampled in Asia (*i.e.,* Russia, China, Mongolia, and Japan, n = 80), and removed variants that had a read depth below five reads per locus, a minor allele frequency less than 0.05, or were present in fewer than 80% of individuals with VCFtools (version 0.1.17) (Danecek et al. 2011). In addition, we removed two individuals with more than 50% missing data and only included one SNP per 10 kb to remove loci that were tightly linked. This pipeline produced 32,029 SNPs derived from 78 individuals across the breeding range of *H. r. rustica* and *H. r. gutturalis* in Asia (Figure 3.1A).

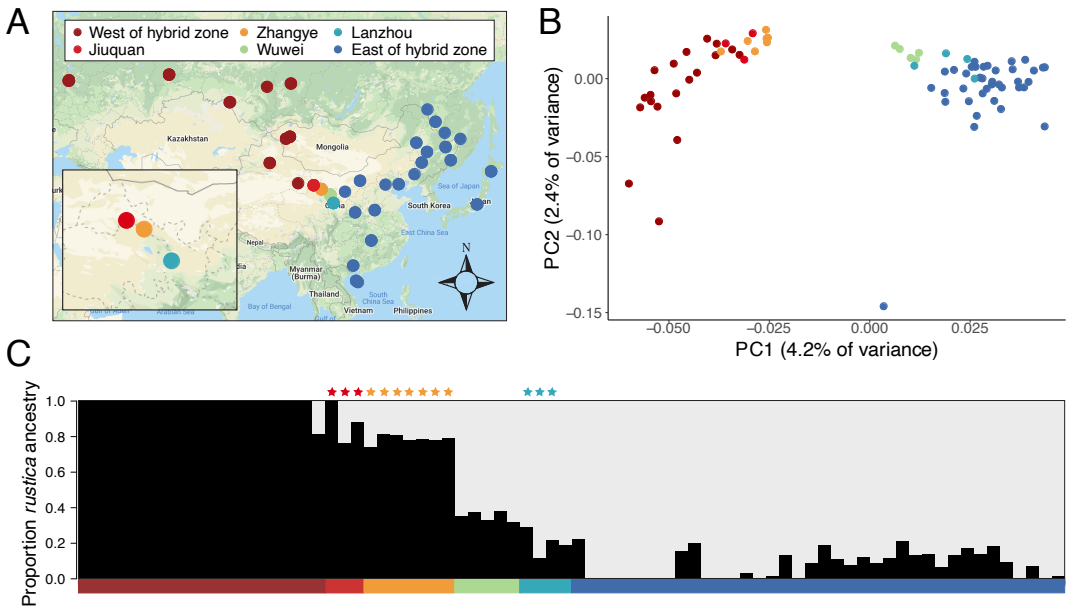


Figure 3.1. (**A**) Sampling locations, (**B**) principal component analysis (PCA) of the genome-wide covariance matrix, and (**C**) admixture proportions generated using whole-genome data for two subspecies of barn swallow (*H. r. rustica* and *H. r. gutturalis*) breeding across Asia (n = 78). Points are colored by sampling location. The inset in (A) shows the three cities where geolocators were deployed. Individuals west of the hybrid zone in China between *H. r. rustica* and *H. r. gutturalis* were sampled from allopatric populations of *H. r. rustica* (dark red), while individuals east of the hybrid zone were sampled from allopatric populations of *H. r. gutturalis* (dark blue). The admixture proportions of the 13 individuals with geolocator tracks are indicated in (C) with stars.

Genome-wide ancestry: Double digest restriction-site associated DNA sequencing

We generated a reduced representation genomic dataset using double digest restriction-site associated DNA sequencing (Peterson et al. 2012), following the protocol in Safran et al. (2016) and Scordato et al. (2017), for 130 individuals that were sampled in July 2016 and May 2017 in the three cities on either side of the hybrid zone in Gansu Province, China (Lanzhou: n = 58, Zhangye: n = 58, Jiuquan: n = 14; Figure 3.1A, Table S3.2). Briefly, we used restriction enzymes *Mse1* and *EcoRI* to fragment loci and ligated sample-specific 8-10 bp barcode sequences for library preparation. Libraries were generated at the University of Colorado Boulder BioFrontiers Institute, size-selected for a 350- to 400-bp region, and sequenced on two replicate Illumina HiSeq 2500 lanes using 100-bp single-end sequencing at the University of Texas Austin Genomic Sequencing and Analysis Facility.

Following sequencing, we detected a lane effect issue stemming from an unintentional difference in fragment size selection from previous libraries. As a result, too few loci were represented by both the current and previous GBS datasets during initial variant calling steps. We were therefore unable to combine data from the newly generated libraries with previously generated GBS data for *H. rustica* from Scordato et al. (2020). In order to leverage data from individuals sampled outside of the putative hybrid zone for ancestry inference, we selected eight individuals that had nearly 100% assignment to *H. r. rustica* or *H. r. gutturalis* from the broader WGS dataset and included these individuals when calling variants for the newly generated ddRAD data.

We used the Stacks (version 2.5) *process_radtags* module (Catchen et al. 2013; Rochette et al. 2019) to demultiplex raw sequencing files and trimmed reads for quality using Trimmomatic (version 0.39) (Bolger et al. 2014), removing bases from read ends with a quality

score < 20 and filtering any reads with an average quality score < 30 or final length < 32 bp. We mapped the filtered ddRAD data, along with WGS data from the eight individuals with high *H. r. rustica* or *H. r. gutturalis* ancestry, to the *H. rustica* reference genome using BWA (version 0.7.17) with program defaults (Li and Durbin 2009). We combined information from all mapping files and called SNPs using BCFtools (version 1.10.2) (Heng et al. 2009). Raw SNP calls were filtered with BCFtools to remove any sites with a median depth among samples < 7. We also used BCFtools to recode any sites with individual depth < 5 or individual genotype quality < 30 as missing data. We used VCFtools to only retain biallelic SNPs that were called in at least 80% of individuals and had a minor allele frequency > 0.05, as well as remove any SNP calls that overlapped with the *H. rustica* genome repeat annotation or were closer than 100 bp in order to reduce the effects of linkage (Danecek et al. 2011). Three samples (two from Zhangye and one from Lanzhou) had > 80% missing data after filtering steps and were therefore removed from analysis. The final filtered VCF table, which included 3,974 SNPs from 127 individuals, was converted to readable input for ancestry inference using Plink (version 1.9) (Purcell et al. 2007).

Population genetic analyses

We used ADMIXTURE (version 1.3) to infer the number of genetic clusters and ancestry proportions of individuals breeding on either side of the hybrid zone in Gansu Province, China relative to allopatric populations of *H. r. rustica* and *H. r. gutturalis* (Alexander et al. 2009). For both the WGS and ddRAD datasets, we ran ADMIXTURE analyses for a series of K values 1-5, using the default cross-validation approach to compare models of K genetic clusters. In both cases, the cross-validation procedure indicated similar support for the K=1 and K=2 models, as these models had a lower cross-validation error than the next best-performing model (K=3).

Because this analysis was designed in part to examine the relative proportion of *H. r. rustica* versus *H. r. gutturalis* ancestry in birds breeding across this region, we interpret ancestry proportions under the K=2 model. A principal component analysis (PCA) of the genome-wide covariance matrix derived from the WGS data using the R function *prcomp* supported the inference of ancestry from two genetic clusters.

To confirm that the ddRAD and WGS-based inferences of ancestry were congruent, we performed a post-hoc comparison of ancestry proportions for the 14 individuals that returned with geolocators, which were sequenced using both WGS and ddRAD approaches. We found a significant positive correlation between the proportion of *H. r. rustica* ancestry inferred using the two datasets (Pearson's r = 0.718, p = 0.004, n = 14; Figure S3.1), indicating that the two methods provide largely consistent estimates of ancestry.

Migratory strategy: Geolocator tracks

During the July 2016 sampling trip, we captured individuals in three cities on either side of the hybrid zone between *H. r. rustica* and *H. r. gutturalis* in Gansu Province, China (Lanzhou: n = 33, Zhangye: n = 33, and Jiuquan: n = 10) and attached Intigeo P55B1-7 geolocators (0.6 g; Migrate Technology, Cambridge, UK) to all captured birds using a Rappole–Tipton leg-loop backpack harness (Rappole and Tipton 1991) made of 0.7 mm elastic cord. The total weight of the tag and harness was less than 5% of average body mass (15.26 ± 1.22 g; mean ± SD) at the time of attachment. In May 2017, we re-sighted 16 of the 76 tagged individuals (21%). However, we only obtained geolocator data from 13 birds (7 males and 6 females), as the battery of one device from Lanzhou stopped working prior to migration and we were unable to capture two individuals. One of the 13 recovered geolocators containing migratory data stopped recording in

March 2017 but was included in subsequent analyses because it collected data throughout the wintering period.

We downloaded light intensity data with IntigeoIF (Migrate Technology) and preprocessed the light readings with the R package *TwGeos* (Wotherspoon et al. 2016), setting a light threshold of 1.5 to automatically determine times of sunrise and sunset. We then visually inspected each twilight period to exclude twilights that exhibited obvious effects of shading. We analyzed the light intensity data with the R package *FLightR* following the workflow detailed in the supplementary material of Rakhimberdiev (2017). Briefly, we derived calibration periods from the data using the *plot_slopes_by_location* function in *FLightR* and optimized the model with one million particles. The *FLightR* model included a spatial mask that prohibited residency over water, allowed individuals to fly a maximum distance of 1500 km between twilights [as in Rakhimberdiev et al. (2016)], and automatically excluded outliers when generating daily estimates of latitude and longitude. We detected stationary periods by running the *stationary.migration.summary* function in *FLightR* with a conservative minimum probability of movement of 0.75 and 5 days as the minimum duration of the stationary period.

To assess the accuracy of the geolocator estimates, we calculated the distance between the true deployment location and the location estimated by the geolocators during the calibration period. In addition, we estimated migratory distance during fall and spring migration by importing the geolocator tracks into QGIS (version 2.18.20) and calculating the length of each track using the length measuring tool (QGIS Development Team 2019). To estimate mean arrival and departure dates from the breeding grounds, we calculated the date by which half of the particles in the *FLightR* model crossed a spatial boundary of one degree longitude west of the breeding site [as in Rakhimberdiev et al. (2016)].

Migratory strategy: Stable isotope analysis

We analyzed the stable-carbon ($\delta^{13}C$) and hydrogen isotope ($\delta^{2}H$) values of the inner-most tail rectrices from the 130 individuals included in the ddRAD dataset to estimate wintering locations for individuals breeding in the three cities on either side of the hybrid zone (*i.e.,* Jiuquan, Zhangye, and Lanzhou). Following the protocol in Scordato et al. (2020), we cleaned tail feathers with a 2:1 mixture of chloroform and methanol to remove oils and surface contaminants, cut 0.5-1 mg from the center of the feather vane, avoiding the tip and base, and rolled the feather into 4 x 6 mm capsules (tin for carbon and silver for hydrogen isotope analysis; Costech Analytical Technologies, Valencia, CA). For carbon isotope measurements, samples were combusted in an elemental analyzer (Carlo Erba NC2500, Milan, Italy) interfaced to an Optima mass spectrometer (VG Micromass, Manchester, United Kingdom) (Fry et al. 1992). Non-exchangeable hydrogen isotope measurements were measured on a thermal conversion elemental analyzer (Thermo Scientific, Bremen, Germany) interfaced to a Delta V mass spectrometer (Thermo Scientific, Bremen, Germany), following conventional methods for carbon reduction and comparative equilibration (Wassenaar and Hobson 2003). All measurements were taken at the U.S. Geological Survey Stable Isotope Laboratory (Denver, CO). We report data in standard delta notation with respect to internationally accepted scales (V-PDB and V-SMOW). Carbon isotope data were normalized to USGS 40 ($\delta^{13}C$ = -26.24 ‰) and 41 ($\delta^{13}C$ = 37.76 ‰), and hydrogen isotope data were normalized to caribou hoof ($\delta^{2}H$ = -157 ‰) and kudu horn ($\delta^{2}H$ = -35.3 ‰). Analytical precision of replicate standards, including two additional secondary standards, was better than ± 0.2‰ for $\delta^{13}C$ and ± 4‰ for $\delta^{2}H$. Two individuals, both sampled in Zhangye, had biologically unrealistic $\delta^{13}C$ and $\delta^{2}H$ values, respectively, and were therefore excluded from subsequent analyses.

We assigned geographic locations of origin for individuals breeding in Jiuquan, Zhangye, and Lanzhou from feather δ^2H values and a geographic model for δ^2H in precipitation (Bowen and Revenaugh 2003) using methods based on Wunder (2010), as the δ^2H isoscape derived from precipitation was more informative in eastern Africa and southern Asia than the $\delta^{13}C$ isoscape for geographic assignments. Feathers collected from birds that returned with geolocators were assumed to be of known origin because tail feathers are grown during the winter and the geolocators provided estimates of a single stationary wintering location per individual. We used these known-origin feathers (n = 13) to calibrate the precipitation model for barn swallow feathers with the function *calRaster* in the R package *assignR* (Ma et al. 2020), and subsequently used the function *pdRaster* to generate a posterior probability of origin raster (*i.e.*, assignment raster) for each of the 116 feathers from birds of unknown wintering location. We limited the spatial extent of the assignment rasters to the extent of the wintering range. We further subdivided the assignment rasters based on recognized Asian flyways for either an African or Asian wintering region (Figure S3.2). Because this subdivision resulted in different numbers of raster cells for each subregion (Africa: n = 20,210 cells; Asia: n = 13,882 cells), we computed the average of the cell values in the assignment raster for each region. If the average for one subregion was more than twice that for the other region, the individual was assigned to the region with the higher average. We assigned individuals as "undetermined" (12/116 individuals; 10%) if there was less than a two-fold difference between the averages.

Statistical analysis

We ran unpaired t-tests to compare migratory distance, timing of migration, WGS-based ancestry estimates, and δ^2H values for individuals with geolocators traveling from separate continents. In

addition, we ran a Kruskal-Wallis test with Dunn's post-hoc test to compare the ddRAD-based ancestry estimates of individuals breeding in the three cities on either side of the hybrid zone. To determine whether ancestry values inferred from ddRAD data were associated with the stable-carbon and stable-hydrogen isotope values of tail feathers, we calculated Pearson's correlation coefficient. Finally, we ran a Fisher's exact test to examine whether wintering continent (Africa vs. Asia) differed by breeding location (eastern vs. western side of the hybrid zone) and a Kruskal-Wallis test with Dunn's post-hoc test to compare distributions of ancestry estimates for individuals that were assigned to Africa, Asia, or unassigned based off of their δ^2H values. All statistical analyses were performed in R version 3.6.2 (R Core Team 2018).

3.4 RESULTS

Genome-wide ancestry differs across the hybrid zone

The PCA of the genome-wide covariance matrix derived from WGS data, which included individuals from the hybrid zone as well as allopatric populations of both subspecies, grouped individuals from Jiuquan and Zhangye with allopatric populations of *H. r. rustica* and individuals from Wuwei and Lanzhou with allopatric populations of *H. r. gutturalis* (Figure 3.1B). These results were further confirmed by the ancestry proportions generated using ADMIXTURE on the WGS data, which indicated that birds sampled in Jiuquan and Zhangye had a large proportion of *H. r. rustica* ancestry (0.81 ± 0.08; mean ± SD), while those sampled in Lanzhou had predominately *H. r. gutturalis* ancestry (0.20 ± 0.07; mean ± SD; Figure 3.1C). In contrast to the PCA, ADMIXTURE detected additional introgression in Wuwei, indicating that individuals breeding near the center of the hybrid zone possess a combination of *H. r. rustica* and *H. r. gutturalis* ancestry (Figure 3.1C).

Ancestry estimates generated using ADMIXTURE on SNPs from the ddRAD dataset, which included individuals sampled in Jiuquan, Zhangye, and Lanzhou, broadly corresponded to the WGS-based estimates (Figure S3.1). The proportion of *H. r. rustica* ancestry significantly differed between individuals breeding on either side of the migratory divide (Kruskal-Wallis: $\chi^2 = 31.55$, df = 2, $p < 0.0001$, n = 127, Figure 3.2). Birds in Jiuquan and Zhangye had a significantly higher proportion of *H. r. rustica* ancestry than individuals breeding in Lanzhou (Dunn's post-hoc test: Jiuquan vs. Lanzhou: $p < 0.0001$, Zhangye vs. Lanzhou: $p < 0.0001$, Jiuquan vs. Zhangye: $p = 0.28$). Relative to the WGS data (Figure 3.1C), the ddRAD data revealed more intermediate individuals and greater variance in ancestry estimates in the area of putative hybridization (Figure 3.2, Figure S3.3); this variation could be due to differences in the hybrid classes of sampled individuals in each dataset, as well as the smaller number of SNPs and lower resolution of the ddRAD data (ddRAD: 3,974 SNPs vs. WGS: 32,029 SNPs).

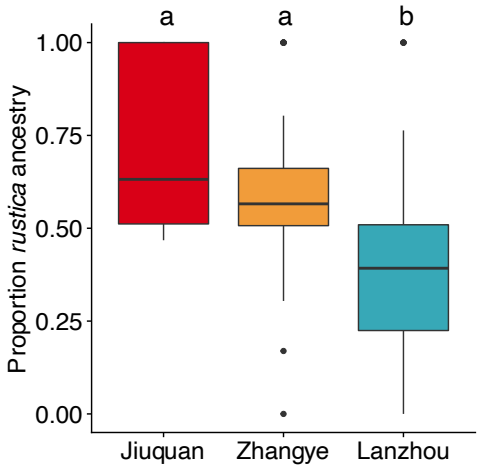


Figure 3.2. Proportion of *H. r. rustica* ancestry derived from ddRAD data for barn swallows (*H. r. rustica* and *H. r. gutturalis*) breeding across a hybrid zone in China (n = 127). Different lowercase letters indicate significant differences between locations (Dunn test with Holm's correction, adjusted $p < 0.05$) and colors correspond to cities on the map in Figure 3.1.

Geolocators document a migratory divide in the hybrid zone

The location estimates derived from the geolocators during the calibration period differed from the true deployment locations in Jiuquan, Zhangye, and Lanzhou by 109 ± 96 km, which is comparable to estimates of similar geolocator studies (Delmore et al. 2012) and smaller than the width of the hybrid zone (620 km from Jiuquan to Lanzhou). The two subspecies exhibited a striking migratory divide that spanned two continents. Individuals breeding on the western side of the hybrid zone (Jiuquan: n = 3 and Zhangye: n = 7) flew north of the Taklamakan Desert and crossed the Arabian Peninsula to overwinter in eastern Africa (Figure 3.3A-B). In contrast, birds breeding on the eastern side (Lanzhou: n = 3) traveled southward across the Qinghai-Tibetan Plateau and overwintered in southern India (Figure 3.3A-B). The geolocator tracks indicated that the two subspecies took divergent migratory routes around the Karakoram Range, a high-altitude mountain range that extends along the borders of India, Pakistan, and China (Figure 3.3A-B). On average, individuals migrating to Africa traveled over twice as far as swallows traveling to India (India: 5,407 ± 834 km vs. Africa: 11,829 ± 1073 km (mean ± SD); unpaired t-test: $p < 0.001$, n = 13, Figure 3.3C).

The WGS ADMIXTURE results revealed that individuals with geolocators that bred in Jiuquan and Zhangye and migrated to Africa had predominately *H. r. rustica* ancestry, while birds that bred in Lanzhou and migrated to India had a much lower proportion of *H. r. rustica* ancestry (Africa: 0.81 ± 0.08 vs. India: 0.21 ± 0.09 (mean ± SD); unpaired t-test: $p = 0.002$, n = 13, Figure 3.3D). In addition, individuals with geolocators that overwintered in Africa had higher δ^2H values relative to birds that overwintered in India (Africa: -19.07 ± 10.44‰ vs. India: -34.59 ± 2.69‰ (mean ± SD); unpaired t-test: $p = 0.002$, n = 13, Figure 3.3E). While uncertainty in location estimates increased around the two equinoxes, and latitudinal estimates were less

accurate than longitudinal estimates, the two subspecies showed clear differences in migratory behavior throughout the annual cycle (Figure S3.4).

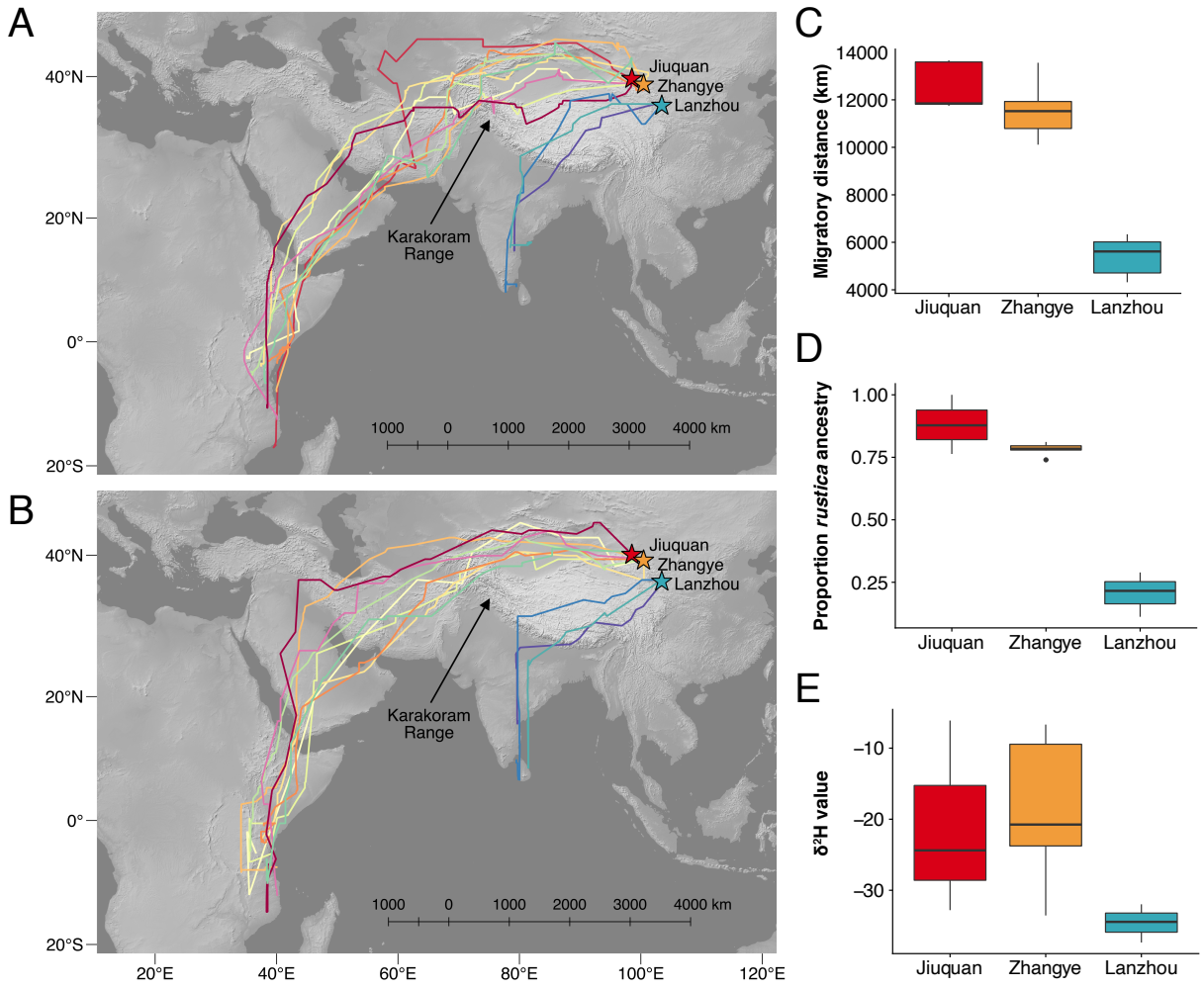


Figure 3.3. Geolocator tracks during **(A)** fall and **(B)** spring migration, **(C)** distance traveled during migration, **(D)** proportion of *H. r. rustica* ancestry generated using whole-genome data, and **(E)** stable-hydrogen (δ^2H) isotope value from tail feathers of barn swallows (*H. r. rustica* and *H. r. gutturalis*) breeding across a hybrid zone in China (n = 13). The tracks were constructed from median geographic position estimates generated twice per day. Individuals that traveled to Africa are shown in warm colors, birds that migrated to India are shown in cool colors, and cities where geolocators were deployed are depicted as stars. Colors correspond to cities on the map in Figure 3.1.

Stable isotope data from a larger sample of individuals provide additional evidence for a migratory divide

The proportion of *H. r. rustica* ancestry for individuals breeding on either side of the hybrid zone (*i.e.,* Jiuquan, Zhangye, and Lanzhou) was positively correlated with the $\delta^{13}C$ ($r = 0.152$, $p = 0.089$, $n = 126$) and $\delta^{2}H$ ($r = 0.290$, $p = 0.001$, $n = 126$) values of their tail feathers, indicating that individuals with predominately *H. r. rustica* ancestry molt their feathers in areas characterized by higher $\delta^{13}C$ and $\delta^{2}H$ values, though the relationship for carbon was only marginally significant (Figure S3.5). Isotope measurements obtained from individuals that were captured in multiple years ($n = 14$) were highly repeatable, with a mean absolute difference between years of 1.06 ± 1.01‰ SD (range = 0.1-3.3‰) for $\delta^{13}C$ and 7.80 ± 5.68‰ SD (range = 1-17‰) for $\delta^{2}H$.

Of the 13 geolocator birds used to calibrate the geographic assignment model, eight of the 10 birds that overwintered in Africa were assigned to Africa and the remaining five individuals, including the three birds that overwintered in India, could not be assigned to either location with a greater than two-fold difference between the average assignment probabilities for the two regions. Of the 116 birds without geolocators breeding in Jiuquan, Zhangye, and Lanzhou, 70 were assigned to Africa, 34 to Asia, and 12 could not be assigned to either location. Wintering location differed by breeding location (*i.e.,* eastern vs. western side of the hybrid zone) for assigned individuals ($p < 0.0001$, $n = 104$, Figure 3.4A). In particular, the odds of an individual breeding on the western side of the hybrid zone (Jiuquan and Zhangye) being assigned to Africa was 148 times greater than the odds of a western-breeding bird being assigned to an Asian wintering locale.

Genome-wide ancestry is associated with migratory behavior

Individuals assigned to different wintering continents on the basis of δ^2H values differed in their ancestry distributions (Kruskal-Wallis: $\chi^2 = 27.32$, df = 2, $p < 0.0001$, n = 116; Figure 3.4B). Pairwise comparisons (Dunn test with Holm's correction) revealed differences between the ancestry estimates of birds assigned to Africa vs. Asia ($p < 0.0001$), and individuals assigned to Africa vs. unassigned ($p = 0.002$), but not between individuals assigned to Asia vs. unassigned ($p = 0.85$; Figure 3.4B). Swallows assigned to Africa had a median ancestry estimate (*i.e.,* proportion *H. r. rustica* ancestry) of 0.57, while those assigned to Asia had a median value of 0.39, which is consistent with the geolocator results that demonstrate that individuals traveling to eastern Africa had predominately *H. r. rustica* ancestry.

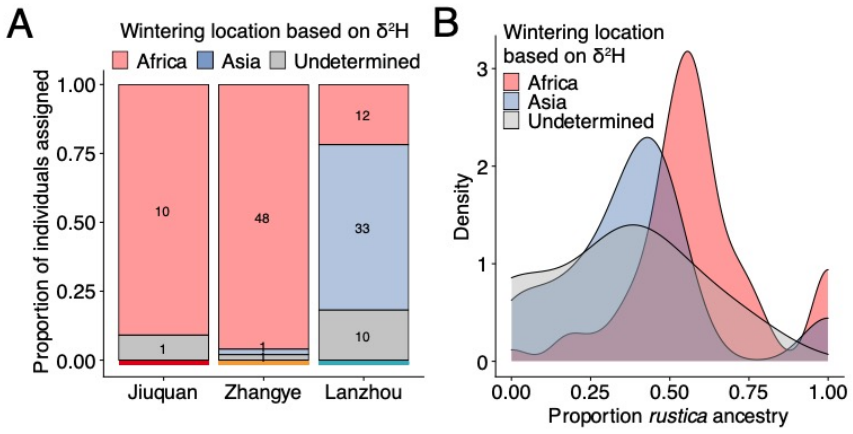


Figure 3.4. (A) Proportion of individuals and **(B)** ancestry distributions of barn swallows (*H. r. rustica* and *H. r. gutturalis*) assigned to Africa (n = 70), Asia (n = 34), or unassigned (n = 12) on the basis of the stable-hydrogen (δ^2H) isotope values of their tail feathers. Colors correspond to cities on the map in Figure 3.1.

Timing of arrival on the breeding grounds differs across the migratory divide

One mechanism by which distinct migratory strategies may contribute to reproductive isolation at migratory divides is divergence in migratory timing and subsequent assortative mating by timing of arrival on the breeding grounds. Barn swallows on either side of the

migratory divide departed from the breeding grounds at similar times (India: September 5 ± 10 days vs. Africa: September 6 ± 6 days (mean ± SD); unpaired t-test: p = 0.83, n = 13). However, swallows that overwintered in India arrived on the breeding grounds almost a month earlier on average than individuals traveling from Africa (India: April 9 ± 6 days vs. Africa: May 8 ± 6 days (mean ± SD); unpaired t-test: p = 0.004, n = 12, Figure 3.5).

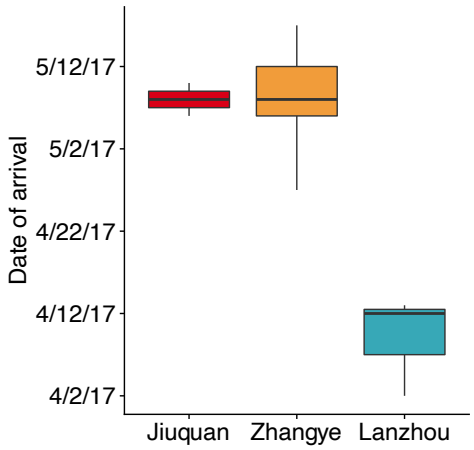


Figure 3.5. Timing of arrival on the breeding grounds derived from geolocators for barn swallows (*H. r. rustica* and *H. r. gutturalis*) breeding across a hybrid zone in China (n = 13). Colors correspond to cities on the map in Figure 3.1.

3.5 DISCUSSION

Divergence in migratory strategy can promote population differentiation and reproductive isolation by exposing organisms to different environments and restricting gene flow between divergent groups. However, few studies have combined emerging animal tracking and genomic sequencing technologies to investigate the role of seasonal migration in the origin and maintenance of species boundaries. Combining data from light-level geolocators, genomic sequencing, and stable isotopes, we documented a striking geographic migratory divide that spans multiple continents and is associated with genomic differentiation across a hybrid zone between barn swallow subspecies. While ongoing gene flow has been reported between *H. r.*

rustica and *H. r. gutturalis* (Scordato et al. 2020), and we detected elevated admixture in Wuwei, located near the center of the hybrid zone, all individuals that returned with geolocators possessed ancestry primarily from one parental subspecies or the other. In addition, the PCA results from the WGS data indicated a lack of hybridization between the subspecies, which suggests fairly strong reproductive isolation and is consistent with the absence of F1 hybrids and steep cline for genomic ancestry previously documented in the hybrid zone (Scordato et al. 2020). Although the lower-resolution ddRAD data revealed greater variance in ancestry estimates in the area of putative hybridization, the ddRAD-based admixture proportions broadly corresponded to the WGS-based estimates, further supporting genomic differentiation across the migratory divide.

In general, barn swallows show remarkably little ecological or genomic divergence across large geographic regions (Scordato et al. 2017). Nonetheless, we found that *H. r. rustica* migrated longer distances from the breeding grounds to overwinter in eastern Africa, while *H. r. gutturalis* traveled shorter distances to southern India. These results were further supported by stable isotope analysis of a larger number of individuals sampled on either end of the hybrid zone. Barn swallows breeding on the western side, with predominately *H. r. rustica* ancestry, had higher stable-carbon and hydrogen isotope values and were assigned to Africa more frequently than those breeding on the eastern side, with *H. r. gutturalis* ancestry. In terrestrial systems, $\delta^{13}C$ values vary systematically with mechanisms of water use efficiency in C3 plants and the proportion of C3 and C4 plants in the environment (Hobbie and Werner 2004). In particular, higher $\delta^{13}C$ values are associated with arid environments dominated by C4 plants, while lower $\delta^{13}C$ values are found in mesic environments with a higher proportion of C3 plants (Dawson et al. 2002; Still et al. 2003). Environmental $\delta^{2}H$ values, in contrast, vary spatially in accordance

with deuterium patterns in local precipitation (Bowen and Revenaugh 2003), which can be exploited for large-scale geographic assignments of migratory species (Hobson 1999; Rubenstein and Hobson 2004). Overall, the observed $\delta^{13}C$ and $\delta^{2}H$ distributions on either end of the hybrid zone, assignment results, and sighting records of the two subspecies (Turner 2010; Sullivan et al. 2014) are consistent with the geolocator data, which indicate that *H. r. rustica* overwinters in eastern Africa, a drier region dominated by more C4 plants, while *H. r. gutturalis* migrates to southern Asia (Still et al. 2003; Scordato et al. 2020).

Differences in migratory distance appeared to affect timing of arrival on the breeding grounds. In particular, individuals on the eastern side of the hybrid zone returned to the breeding grounds over three weeks earlier than individuals on the western side, likely due to their shorter migratory route. Divergence in timing of arrival is a widespread pattern at migratory divides, where populations on either side often migrate different distances to their non-breeding grounds (Bensch et al. 1999; Ruegg et al. 2012). For example, since the evolution of a novel migratory route in Eurasian blackcaps (*Sylvia atricapilla*), northwest migrants that overwinter in the United Kingdom arrive on sympatric breeding grounds in Germany significantly earlier on average than blackcaps that migrate along their historical route to the Mediterranean (Bearhop et al. 2005; Rolshausen et al. 2010). In addition, in a well-studied hybrid zone between subspecies of Swainson's thrush (*Catharus ustulatus*) in British Columbia, the short-distance coastal migrant (*C. u. ustulatus*) arrives earlier on the breeding grounds than the longer-distance inland migrant (*C. u. swainsoni*) (Ruegg et al. 2012).

Differential timing of migration as a result of divergent migratory strategies can promote reproductive isolation if individuals mate assortatively by timing of arrival on the breeding grounds (Turbek et al. 2018). Eurasian blackcaps that follow the novel northwestern migratory

route and arrive earlier on the breeding grounds are more likely to pair with other northwestern migrants than southwestern migrants that overwinter in the Mediterranean (Bearhop et al. 2005). Since the 1960s, when the new migratory route in blackcaps was first documented, genetic and phenotypic differences have accumulated between the two migratory forms (Rolshausen et al. 2009), supporting the idea that pre-mating isolation due to divergence in arrival times can drive population differentiation.

While we found that divergent migratory behavior corresponds to differences in genome-wide ancestry in barn swallows, we were unable to analyze whether assortative mating by migratory timing limits gene flow between the subspecies, as geolocators were only deployed on either side of the migratory divide. However, parental individuals of *H. r. rustica* and *H. r. gutturalis* do not come into direct contact in the center of the hybrid zone, which consists solely of admixed individuals, and previous work within the hybrid zone found little support for assortative mating by carbon isotope value (Scordato et al. 2020). Further population-level studies from the hybrid zone center are needed to determine whether the lack of overlap in arrival times between individuals on either side of the migratory divide contributes to the maintenance of subspecies boundaries.

In addition to assortative mating by timing of arrival, divergent migratory behavior could also prevent gene flow through post-mating isolation. For example, hybrids at migratory divides may inherit maladaptive trait combinations or intermediate migratory routes that expose them to increased mortality during migration (Rohwer and Irwin 2011; Delmore and Irwin 2014). Migratory orientation is at least partially genetically determined in songbirds (Berthold 1991; Liedvogel et al. 2011), and several studies have identified genomic regions associated with divergent migratory phenotypes between closely related populations (Delmore et al. 2016;

Lundberg et al. 2017). In addition, experimental crosses in Eurasian blackcaps, as well as natural hybridization in Swainson's thrushes, have revealed that hybrid individuals can inherit intermediate migratory directions relative to parental populations (Helbig 1991; Delmore and Irwin 2014; Delmore et al. 2016).

Uncertainty in latitudinal estimates derived from the geolocators increased around the equinoxes, when birds migrate to and from the wintering grounds. However, the geolocator tracks clearly indicated that the two subspecies took divergent migratory routes around the Karakoram Range, which contains eight peaks over 7,500 meters and constitutes the second highest mountain range in the world. Furthermore, all of the barn swallows that returned with geolocators had a large proportion of ancestry from one parental subspecies or the other, and the migratory routes of hybrids remain unknown. Barn swallows have been observed above 3,000 m at various locations along the Qinghai-Tibetan Plateau during the summer (Farrington 2016). However, if hybrid individuals within the contact zone attempt to directly cross arid and mountainous regions that offer few opportunities for refueling, elevated mortality during migration could lead to post-mating isolation and maintain genetic differentiation between the subspecies.

While several geographic features in central Asia, particularly the Qinghai-Tibetan Plateau, have been proposed as major barriers to migration, no previous study has directly tracked Asian passerines to examine how small-bodied birds navigate these inhospitable barriers. We document a migratory divide characterized by divergent routes around the high-altitude Karakoram Range that is closely associated with genomic differentiation in barn swallows. These findings are consistent with the long-standing hypotheses that (1) divergent migratory behavior can prevent gene flow and facilitate divergence and (2) inhospitable geographic

features may have contributed to the diversification of Asian songbirds by influencing migratory patterns.

CHAPTER 4

RAPID SPECIATION VIA THE EVOLUTION OF PRE-MATING ISOLATION IN THE IBERÁ SEEDEATER[3]

4.1 ABSTRACT

Behavioral isolation can catalyze speciation and permit the slow accumulation of additional reproductive barriers between co-occurring organisms. We illustrate how this process occurs by examining the genomic and behavioral bases of pre-mating isolation between two bird species (*Sporophila hypoxantha* and the recently discovered *S. iberaensis*) that belong to the southern capuchino seedeaters, a recent, rapid radiation characterized by variation in male plumage coloration and song. Though these two species co-occur without obvious ecological barriers to reproduction, we document behaviorally-mediated species recognition and strong assortative mating associated with genomic regions underlying male plumage patterning. Plumage differentiation likely originated through the reassembly of standing genetic variation, indicating how novel sexual signals may quickly arise and maintain species boundaries.

4.2 INTRODUCTION

Organisms in the early stages of speciation provide an opportunity to understand the processes that govern reproductive isolation between taxa (Butlin et al. 2012). Pre-mating isolation (*e.g.,* ecological or behavioral mechanisms that prevent individuals from interbreeding) is a powerful barrier that can separate sympatric species early in divergence (Sobel and Streisfeld

[3] This chapter was published in *Science* with Melanie Browne, Adrián S. Di Giacomo, Cecilia Kopuchian, Wesley M. Hochachka, Cecilia Estalles, Darío A. Lijtmaer, Pablo L. Tubaro, Luís Fábio Silveira, Irby J. Lovette, Rebecca J. Safran, Scott A. Taylor, and Leonardo Campagna

2015; Lackey and Boughman 2017; Uy et al. 2018). While post-mating barriers, such as genetic incompatibilities, take longer to accumulate than the time to speciation of many taxa (Coyne and Orr 1989; Price and Bouvier 2002), learned or genetic preferences can diverge over shorter timescales and generate assortative mating (Seehausen et al. 2008; Grant and Grant 2018; Xu and Shaw 2019), fueling rapid speciation and paving the way for the accumulation of additional reproductive barriers (Coyne and Orr 1989; Mendelson 2003; Lackey and Boughman 2017). Tracking mating decisions among wild populations early in speciation can improve our understanding of how behavioral isolation promotes divergence.

Southern capuchino seedeaters (*Sporophila*) are one of the most rapid avian radiations, showing remarkably low levels of ecological and genomic divergence (Campagna et al. 2012, 2017). Like Lake Victoria cichlids, where differences in male coloration promoted rapid diversification (Selz et al. 2014), the Neotropical southern capuchinos radiated within the last million years to form ten predominantly sympatric species that differ primarily in male plumage coloration and song (Campagna et al. 2012, 2017). Field experiments suggest that divergent male traits govern conspecific recognition and territorial defense (Benites et al. 2015). Nonetheless, viable hybrids between capuchino species are readily produced in the field (Medolago et al. 2020) and captivity (Campagna et al. 2018), suggesting a lack of genetic incompatibilities.

Here we take advantage of the identification of *S. iberaensis* (the Iberá Seedeater), a newly described species from Iberá National Park, Argentina, where six other southern capuchinos co-occur during the breeding season (BirdLife International 2020), to study the importance of pre-mating barriers early in speciation. *Sporophila iberaensis* was first observed in October of 2001 (Di Giacomo and Kopuchian 2016), has a breeding range contained entirely within that of *S. hypoxantha* (Figure 4.1A-B, Figure S4.1), and breeds primarily in the northern

portion of the Iberá wetlands (in the 111,000 hectare San Nicolás Reserve), where both species hold neighboring territories. Unlike its congeners, *S. iberaensis* is increasing in local abundance (Figure 4.1C, Figure S4.1C). The species' small breeding range, combined with the fact that this region was unexplored from an ornithological perspective until the beginning of the 21st century due to a lack of public roads, suggests that *S. iberaensis* likely already existed in the area and went undescribed (Supplementary Text). This is consistent with other southern capuchino species that have small and restricted breeding ranges (*e.g., S. melanogaster* and *S. nigrorufa* (Figure S4.1A) (BirdLife International 2020)) and another more distantly related species in this taxonomically challenging genus (*S. beltoni*), which has a limited breeding range and was only recently identified in South America (Repenning and Fontana 2013).

Throughout two breeding seasons, we located and monitored 128 nests of *S. hypoxantha* and *S. iberaensis*, the only two southern capuchinos observed successfully breeding in the San Nicolás Reserve in Iberá National Park during the study (Figure 4.1D-E; Table S4.1). We collected samples for genomic analyses from 80 nestlings and 126 adults and performed behavioral experiments of these two species to examine 1) the role of assortative mating in the maintenance of species boundaries, 2) the phenotypic traits underlying species recognition, 3) the genomic basis of such traits, and 4) the origin of these genomic variants.

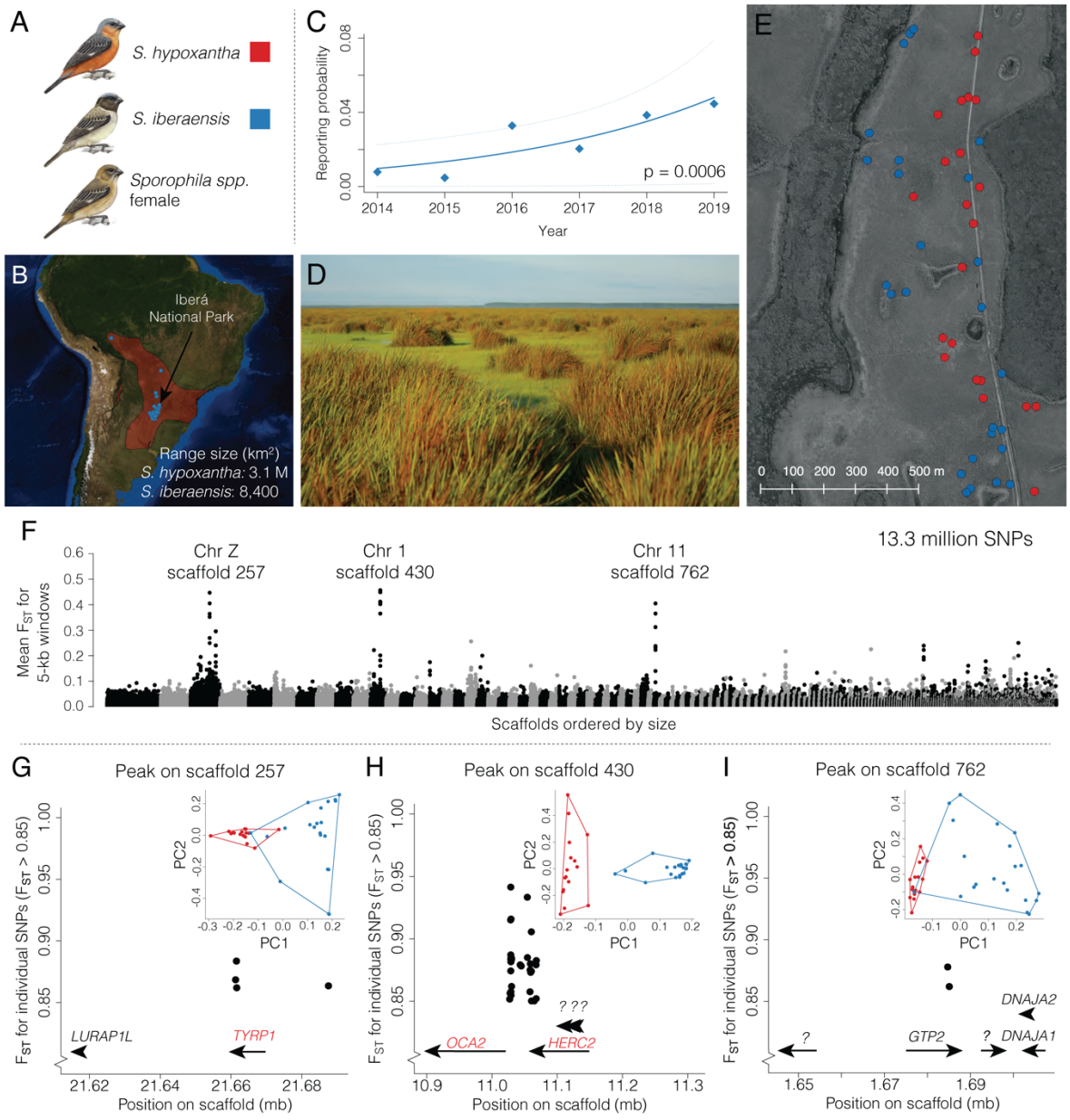


Figure 4.1. Geographic context and genomic characterization of the study species. **(A)** Plumage phenotype and **(B)** breeding distribution of *S. hypoxantha* (red) (BirdLife International 2020) and *S. iberaensis* (blue circles are observations in the eBird database). The arrow indicates the study site location. **(C)** Increase in reporting rate probability for *S. iberaensis* in the eBird database (dotted lines are 95% prediction intervals for the estimated probabilities). **(D)** Typical breeding habitat. **(E)** Spatial distribution of 49/128 nests of *S. hypoxantha* and *S. iberaensis* found during the study. **(F)** Pattern of genomic differentiation between individuals of *S. hypoxantha* ($N = 16$) and *S. iberaensis* ($N = 21$). Divergence peaks are labeled according to their scaffold and corresponding chromosome in the zebra finch assembly. The plot contains the 733 largest scaffolds. **(G-I)** Genomic locations of individual SNPs with $F_{ST} > 0.85$ on scaffold **(G)** 257, **(H)** 430, and **(I)** 762. Genes within 50-kb of these SNPs are depicted with arrows drawn to scale,

with those involved in coloration highlighted in red. The insets show PCAs of the SNPs under the peaks.

4.3 RESULTS

S. hypoxantha and S. iberaensis show low genomic differentiation

We examined the degree of genomic divergence between these species using shotgun short-read whole-genome sequences from 16 individuals of *S. hypoxantha* and 21 individuals of *S. iberaensis* (20 males and 17 females; Table S4.2), identifying ~13.3 million single nucleotide polymorphisms (SNPs). The 42 SNPs showing the highest differentiation ($F_{ST} > 0.85$, max = 0.94) were concentrated in three relatively narrow (30-50 kb) divergence peaks, which were located on separate chromosomes (1, 11, and sex chromosome Z; Figure 4.1F) and exhibited increased absolute sequence divergence (D_{XY}; Figure S4.2). As among other capuchinos (Campagna et al. 2017), *S. hypoxantha* and *S. iberaensis* are characterized by extremely low genomic differentiation (mean $F_{ST} = 0.006 \pm 0.059$ SD) and no mitochondrial divergence (Figure S4.3). However, individuals clustered by species in a genome-wide principal component analysis (PCA; Figure S4.4). Individuals also clustered by species in separate PCAs performed with the SNPs from each peak (Figure 4.1G-I), but only one peak completely differentiated the two species (scaffold 430; Figure 4.1H). Despite being located on different chromosomes, the three regions showed high values of linkage disequilibrium within and among the peaks (Figure S4.5), indicating their co-inheritance. The lack of fixed differences (*i.e.,* $F_{ST} = 1$) between *S. iberaensis* and *S. hypoxantha* among the 42 highly differentiated SNPs identified in our genome-wide F_{ST} analysis motivated us to search for the extent of shared variants between the species in the divergence peaks. First, we PCR-amplified and Sanger sequenced a ~700 bp region that included 15 of the 64 SNPs with F_{ST} greater than 0.79 within the peak on scaffold 430 (mean $F_{ST} = 0.872$

± 0.028 SD), allowing us to assess the genetic variation in this region for a sample of 202 individuals. We observed 21 out of 200 *S. iberaensis* haplotypes that grouped with those of *S. hypoxantha* (Figure 4.2). While each species had a common haplotype, a few *S. iberaensis* individuals carried the *S. hypoxantha* haplotype, and a small proportion of haplotypes appeared to be intermediate. Some of the intermediate haplotypes belonged to *S. iberaensis* individuals and clustered with *S. hypoxantha*, yet we also observed intermediate haplotypes in *S. hypoxantha* that did not cluster with *S. iberaensis* (Figure 4.2). We obtained similar results when conducting this haplotype-based analysis on the 37 individuals with whole-genome sequencing data for all the variants found in the peak on scaffold 430 (Figure S4.6) and the SNPs showing the highest level of differentiation within the peaks on scaffolds 430 (Figure S4.7) and 257 (Figure S4.8, Supplementary Text). Taken together, these findings are consistent with *S. hypoxantha* variants segregating within *S. iberaensis* at the sites showing the highest differentiation between both species (and to a lesser extent in the reverse direction), either due to incomplete lineage sorting or past events of hybridization.

Divergence peaks contain plumage coloration genes

We identified 12 genes within the divergence peaks (Figure 4.1G-I; Table 4.1). Two peaks (scaffolds 257 and 430) contained genes known to be involved in melanic coloration (*TYRP1*, *OCA2*, and *HERC2*; Figure 4.1G-I; Table 4.1) (Abolins-Abols et al. 2018). Most highly differentiated SNPs (98%) were located in non-coding regions (Table S4.3), which may contain cis-regulatory elements that generate phenotypic variation (Campagna et al. 2017). Although genes of small effect located outside of the divergence peaks could contribute to phenotypic

differentiation, only 1.4% of SNPs in the genome had moderate F_{ST} values ($F_{ST} > 0.2$; Figure S4.9), suggesting that high differentiation is largely confined to the F_{ST} peaks.

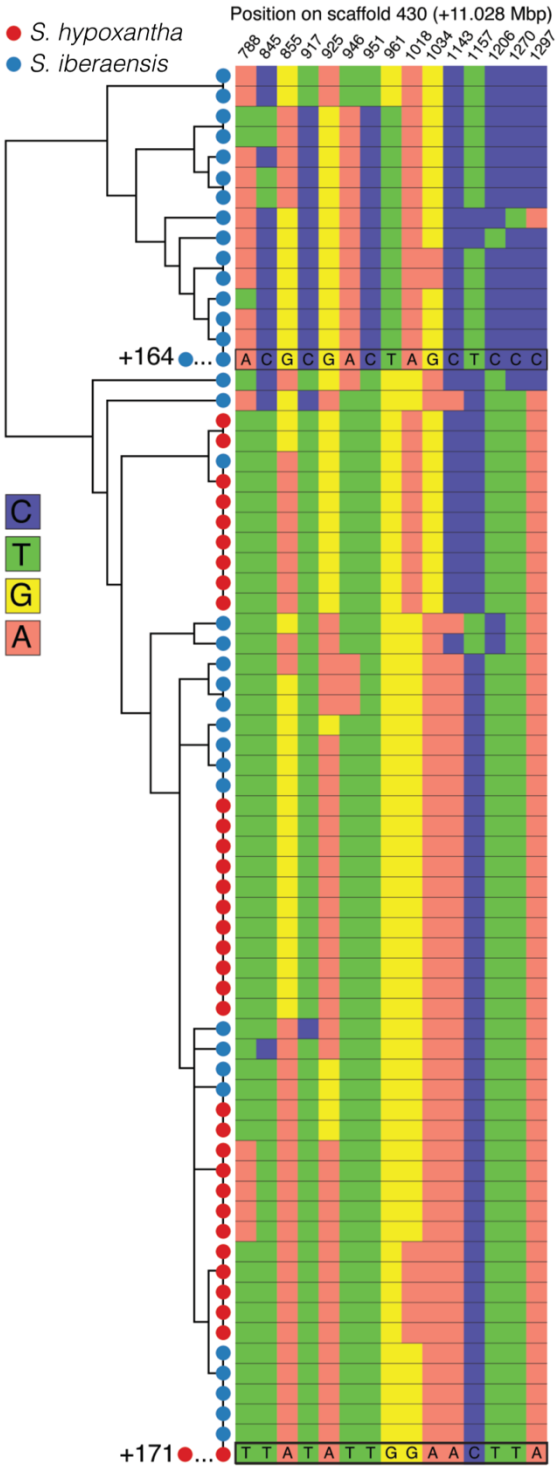


Figure 4.2. Clustering of haplotypes obtained from the region of highest differentiation on scaffold 430. Phased genotypes of males, females, and nestlings of *S. hypoxantha* and *S. iberaensis* ($N = 202$) for 15 highly divergent SNPs located in the peak on scaffold 430 generated from either whole-genome or Sanger sequence data (~700 bp). Each row represents a single chromosome, and each individual is represented twice in the tree. The four nucleotides are color-coded as indicated in the upper left corner. *S. hypoxantha* individuals and the majority of *S. iberaensis* birds have species-specific haplotypes. However, 17/100 (17%) *S. iberaensis* birds possessed one haplotype that clustered with *S. hypoxantha* and two *S. iberaensis* individuals (2%) clustered with *S. hypoxantha* on the basis of both haplotypes. The most common haplotype for each species is indicated at the bottom of the two main clusters. For graphical clarity, identical copies of each of these common haplotypes were omitted from the tree.

Table 4.1. Regions of elevated genomic differentiation between the two species.

Scaffold	Chr	Peak size (kb)	Highest F_{ST} (over 5-kb window)	No. SNPs with F_{ST} > 0.85	No. genes (known function)	Coloration genes	Coloration gene function
257	Z	50	0.533	4	2 (2)	*TYRP1*	Enzyme involved in the production of melanin (Lyons et al. 2005; Kenny et al. 2012; Domyan et al. 2014; Li et al. 2019)
430	1	45	0.485	36	5 (2)	*OCA2, HERC2*	*OCA2*: Melanosomal transmembrane protein (Sturm and Frudakis 2004; Visser et al. 2014; Klaassen et al. 2018), *HERC2*: Contains a regulatory sequence that controls *OCA2* expression (Visser et al. 2012)
762	11	30	0.435	2	5 (3)	-	-

S. hypoxantha and S. iberaensis mate assortatively in sympatry

S. hypoxantha and *S. iberaensis* females do not show clear morphological characters that allow their identification to species. However, given that birds can detect wavelengths in the ultraviolet range (300-400 nm) that are not perceived by humans (Cuthill et al. 2000), we examined the extent to which females of *S. hypoxantha* and *S. iberaensis* overlap in plumage coloration from an avian visual perspective. Benites *et al.* (2010) detected coloration differences among females of four capuchino species that could be perceived by birds; however, we found

that a large percentage of the convex hulls encompassing females of *S. iberaensis* in tetrahedral color space (a model of avian vision) were contained within those of *S. hypoxantha*, and the species largely overlapped in coloration across the avian visual spectrum (Figure 4.3). Therefore, we used the divergent genomic regions to identify females to species and quantify assortative mating. Paired males and females (N = 17 pairs) clustered together in a tree based on whole-genome data (Figure 4.4A) and shared the same F_{ST} peaks (Figure S4.10), indicating a lack of hybrid pairs. We expanded this analysis to pairs for which we lacked whole-genome data (N = 23) by using double-digest restriction site-associated DNA (ddRAD) sequencing to genotype all sampled males, females, and nestlings (N = 206) at ~61.5 thousand SNPs. Despite the extremely shallow genomic differentiation between these species, individuals clustered into two groups in a PCA, matching the phenotype of the male attending each nest (Figure 4.4B). This signal was derived mainly from the cumulative effect of SNPs with low F_{ST} values, as the ddRAD data only contained 28 SNPs that fell within the F_{ST} peaks identified from the whole-genome data (Figure S4.11A) and showed the same pattern when those SNPs were excluded from the PCA (Figure S4.11B).

Because mating outside of the social pair bond is common in birds (Griffith et al. 2002), we also used 281 highly informative ddRAD loci to evaluate patterns of paternity. While the rate of extra-pair mating was very high (> 52%; 35/67 offspring with known social fathers), all extra-pair offspring that matched candidate fathers in the dataset were sired by males of the same species as their social father (N = 18; Table S4.4). In addition, both social (N = 40) and genetic pairs (N = 27) clustered by species based on genomic PC1 score (Figure 4.4C), indicating that assortative mating is maintained via both social and extra-pair mating.

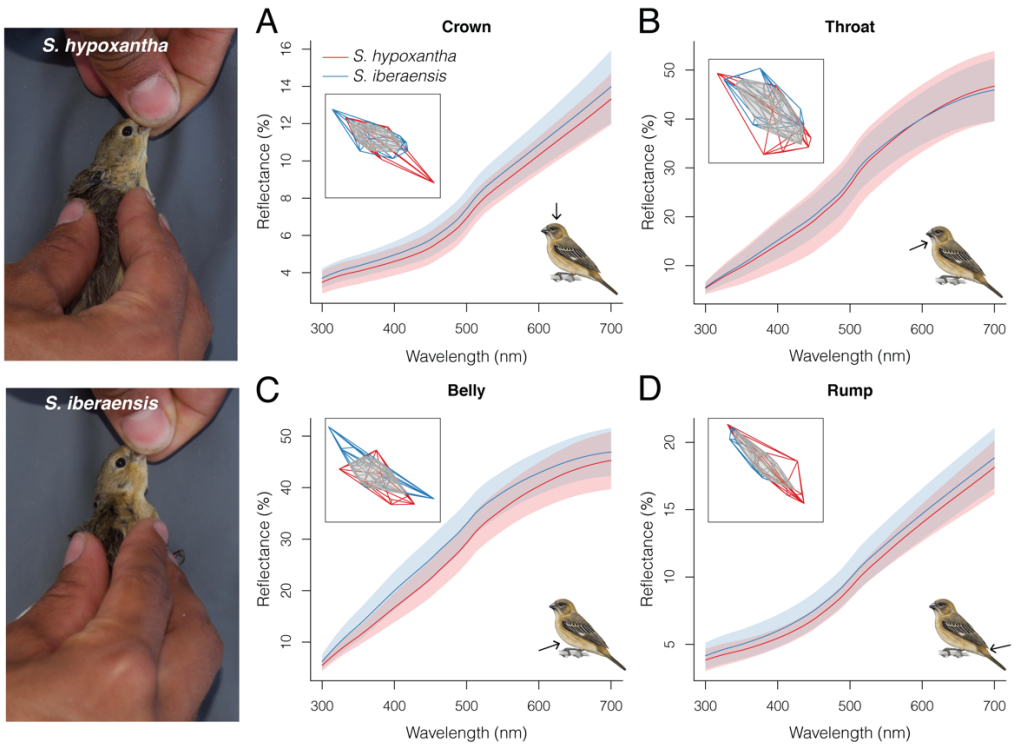


Figure 4.3. Females of *S. hypoxantha* and *S. iberaensis* overlap in plumage coloration. Phenotype, degree of overlap, and reflectance patterns across the avian visual spectrum for the **(A)** crown, **(B)** throat, **(C)** belly, and **(D)** rump of *S. hypoxantha* ($N = 22$) and *S. iberaensis* ($N = 20$) females. The lines indicate mean reflectance for each group, the shaded areas depict the standard deviation, and the arrows indicate the location of each measured plumage patch. The gray polygons in the insets show the extent of overlap in tetrahedral color space between the two species for the crown (43%), throat (72%), belly (58%) and rump (83%).

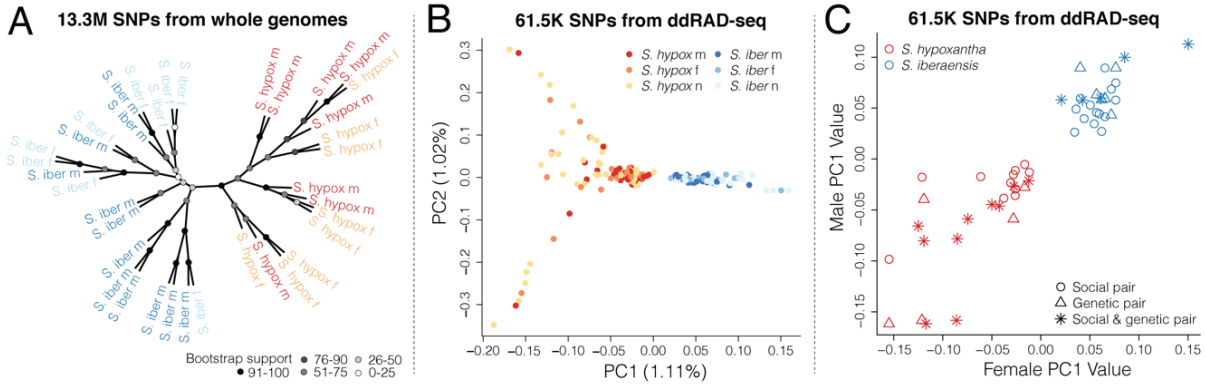


Figure 4.4. No evidence of hybridization through social or extra-pair mating. **(A)** Whole-genome coalescent tree showing the relationship between males and females of *S. hypoxantha* and *S. iberaensis* ($N = 37$). **(B)** PCA from double-digest restriction site-associated DNA (ddRAD) sequencing data depicting genomic differentiation between males, females, and

nestlings of *S. hypoxantha* and *S. iberaensis* (*N* = 206). **(C)** Genomic PC1 scores of males and females of *S. hypoxantha* and *S. iberaensis* for social (*N* = 40) and genetic pairs (*N* = 27; *i.e.,* pairs that fertilized within-pair or extra-pair offspring). For all plots, females were classified based on the phenotype of their social mate.

Species discrimination is based on plumage and song traits

Sporophila hypoxantha and *S. iberaensis* mate assortatively despite holding neighboring territories during the breeding season (Figure 4.1E), breeding synchronously (Figure S4.12), and foraging together on the same grasses (Turbek et al. 2019). In addition to male plumage patterning, capuchinos differ in song (Figure S4.13), a culturally transmitted trait acquired primarily through social learning in songbirds, though there is a genetic component of early song discrimination (Wheatcroft and Qvarnström 2017). Therefore, differences in male plumage patterning and song, rather than temporal or spatial barriers to reproduction, likely mediate mate choice and prevent interbreeding through genetic and/or imprinting mechanisms (Benites et al. 2015; Grant and Grant 2018; Uy et al. 2018).

To test the roles of divergent plumage patterning and song in species recognition and pre-mating isolation, we presented territorial males of *S. hypoxantha* (*N* = 40) and *S. iberaensis* (*N* = 36) with all combinations of conspecific and heterospecific capuchino song and plumage (using song playback and artificial mounts, see Figure 4.5A-C, Figure S4.14), as well as that of a sympatric and ecologically similar heterospecific control (*S. collaris*), and assessed their behavioral responses. Across 240 trials (24 per treatment/species), we recorded aggressive behaviors and generated a response intensity score using PCA (Figure S4.15). Each species responded most aggressively to the combination of conspecific song and plumage, exhibited intermediate responses to the treatments with mismatched traits, and largely ignored the heterospecific capuchino traits and those of the control species (Figure 4.5D-E, Figure S4.16, Table S4.5). Generalized linear mixed models confirmed that both song and plumage are used to

recognize sexual competitors, with significant effects on the intensity of the males' response (song/plumage: $P < 0.0001$) and attack behavior (song: $P = 0.005$, plumage: $P = 0.012$) in both species (Table 4.2).

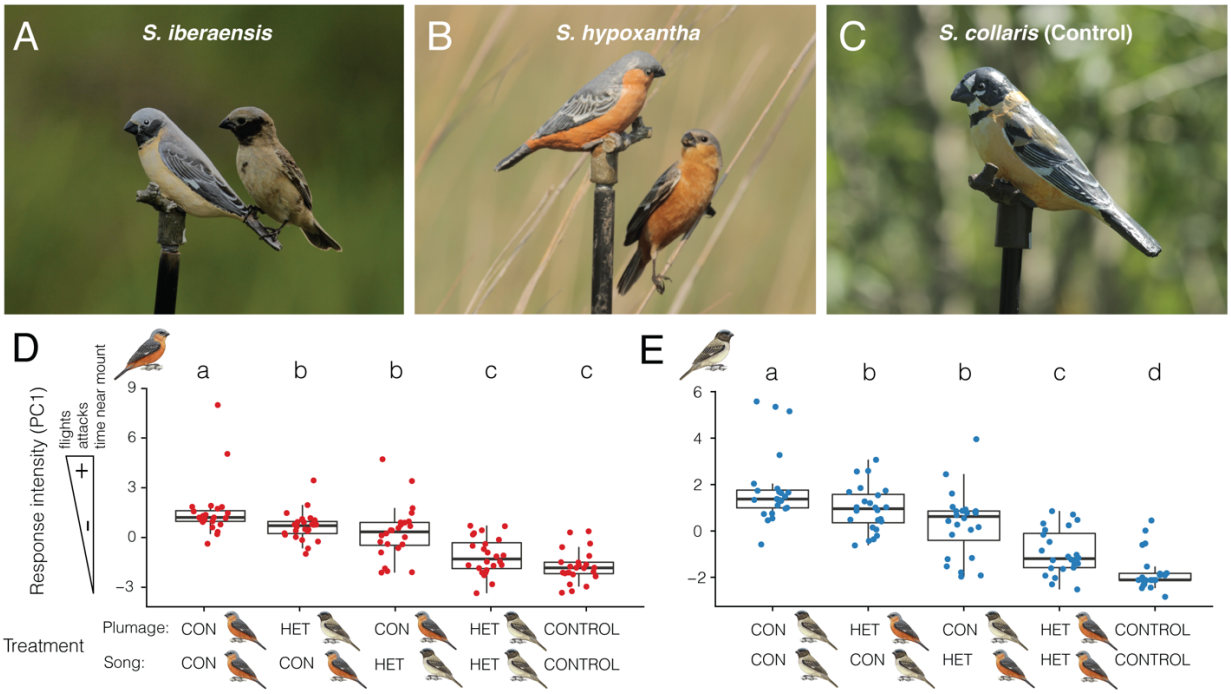


Figure 4.5. Territorial males of both species respond most aggressively to conspecific song and plumage. Artificial mounts of **(A)** *S. iberaensis*, **(B)** *S. hypoxantha*, and **(C)** *S. collaris* (the control) alongside breeding males of the two capuchino species. Two mounts were created per species for use in the behavioral experiment. **(D-E)** Behavioral response intensity (PC1) of territorial males of *S. hypoxantha* and *S. iberaensis*, respectively, to combinations of conspecific (CON), heterospecific capuchino (HET), and control (CONTROL) song and plumage. Different letters indicate statistical significance between treatment groups (Tukey HSD; adjusted $P < 0.05$, $N = 120$ per species).

Table 4.2. Species discrimination is based on both plumage and song. Generalized linear mixed model results examining the behavioral responses of territorial males of *S. hypoxantha* and *S. iberaensis* to mount presentations and song playbacks when the heterospecific control (*S. collaris*) trials were excluded. Significant results ($P < 0.05$) are highlighted in bold. Plumage (conspecific vs. heterospecific) and song (conspecific vs. heterospecific) had a significant effect on response intensity regardless of whether outliers (observations outside 1.5 * interquartile range) in each treatment group were included ($N = 192$) or excluded ($N = 179$; plumage: $P < 0.0001$, song: $P < 0.0001$), while the species of the focal male did not affect behavioral response. We detected an additional significant interaction between song and plumage on response

intensity ($P = 0.03$) when outliers were removed, which could indicate a synergistic effect when both traits belong to the same species.

Response intensity (PC1)[1]	**Estimate**	**Std. error**	***t* value**	***P* value**
Intercept	-1.63	0.29	-5.70	**0.0007**
Species	-0.16	0.26	-0.62	0.535
Plumage	0.93	0.18	5.04	**<0.0001**
Song	1.41	0.19	7.58	**<0.0001**
Plumage × Song	0.24	0.26	0.92	0.360

Attack behavior[2]	**Estimate**	**Std. error**	***z* value**	***P* value**
Intercept	0.11	0.78	0.14	0.889
Species	1.19	1.00	1.20	0.232
Plumage	-2.76	1.10	-2.52	**0.012**
Song	-3.19	1.14	-2.79	**0.005**
Plumage × Song	-20.38	12802.17	-0.002	0.999

[1]Model included male ID (SD = 0.89; 95% CI of SD = 0.70-1.12) and female presence (SD = 0.26; 95% CI of SD = 0.03 – 1.21) as random effects.

[2]Model included male ID (SD = 2.90; 95% CI of SD = 1.34-6.28) as a random effect.

Existing mutations in novel combinations underlie the plumage phenotype of S. iberaensis

To investigate the origin of the novel *S. iberaensis* plumage phenotype, we examined genomic differentiation across the broader capuchino radiation (~28.2 million SNPs across 127 individuals from 12 species). We generated phylogenies using maximum likelihood for the entire genome and the regions containing divergence peaks. The whole-genome tree showed patterns consistent with recent speciation (Figure 4.6A, Figure S4.17), such as a lack of species-level monophyly possibly due to hybridization and incomplete lineage sorting, a result that was further supported by demographic modeling (Supplementary Text, Figure S4.18). Despite this phylogenetic uncertainty, *S. iberaensis* formed a clade, as did most individuals from other species with restricted ranges (see *S. melanogaster* in green and *S. nigrorufa* in yellow; Figure 4.6A, Figure S4.17). In contrast, *S. iberaensis* did not form a species-specific clade in the phylogenies derived from the regions containing divergence peaks (see arrows in Figure 4.6B-C, Figure S4.19), unlike most *S. melanogaster* (indicated with a circle in Figure 4.6B) and *S.*

ruficollis (indicated with a circle in Figure 4.6C) individuals. Although multiple species shared variants with *S. iberaensis* at the individual divergence peaks (*e.g., S. ruficollis* in peak 257 and five other species in peak 430; Figure 4.6B-C, Figure S4.19), the particular combination found only in *S. iberaensis* distinguished it from other capuchinos (note that no other capuchino shares variants with *S. iberaensis* at both divergence peaks; Figure 4.6B-C, Figure S4.20). This result implies that the *S. iberaensis* phenotype likely arose through the reshuffling of standing genetic variation that already existed within the other southern capuchinos, providing a mechanism for rapid speciation without the long period required for relevant mutations to arise *de novo* (Meier et al. 2018; Marques et al. 2019).

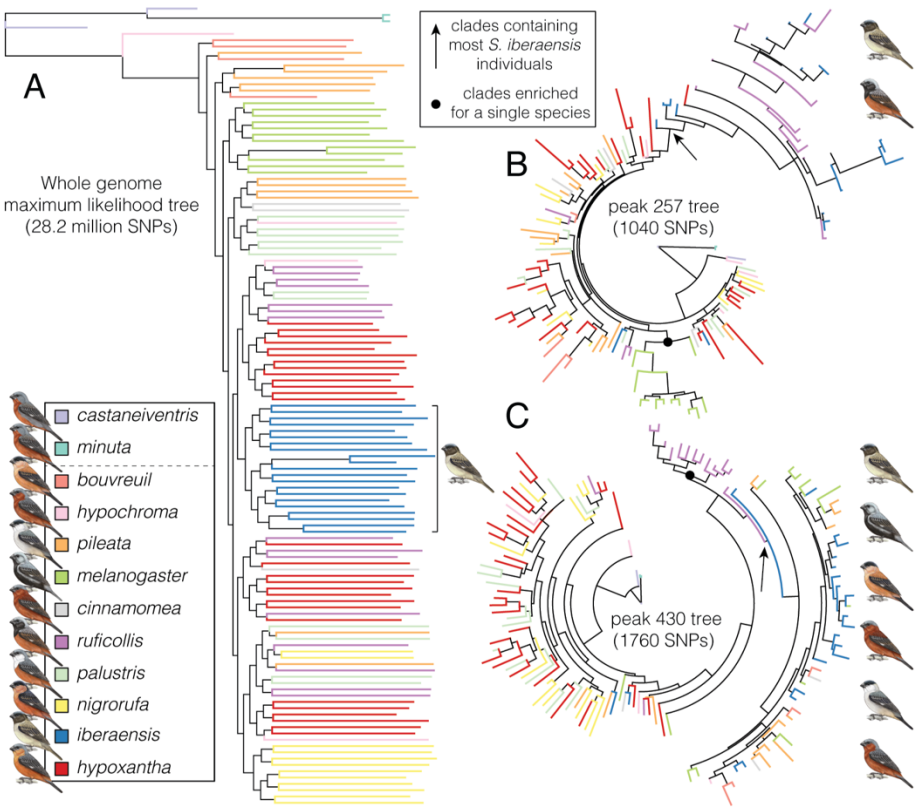


Figure 4.6. *S. iberaensis* is monophyletic but shares variants at divergence peaks with other capuchinos. Capuchino phylogeny inferred using maximum likelihood based on **(A)** whole-genome data and SNPs from the peaks on **(B)** scaffold 257 and **(C)** scaffold 430. The black square bracket and arrows indicate clades containing all or most *S. iberaensis* individuals and the

black circles indicate clades of other species with species-specific variants at the peaks of differentiation. Outgroups are shown above the dashed line in the legend.

4.4 DISCUSSION

Our findings point to pre-mating isolation through assortative mate choice, based on genetically inherited (plumage color) and culturally inherited traits (song, but see Wheatcroft and Qvarnström (2017)), as a primary mechanism promoting divergence between these co-occurring capuchino species. While we never observed hybrid pairs during this study, selection against intermediate traits (plumage patterns or songs) or mismatched plumage and song traits in hybrids could further strengthen assortative mating through reinforcement (Howard 1993). Most divergence peaks in capuchinos (Campagna et al. 2017) and one of the three peaks between *S. iberaensis* and *S. hypoxantha* are located on the sex chromosome Z. Loci on sex chromosomes are thought to have a disproportionate effect on hybrid fitness (large-Z effect (Ellegren 2009)) and may have played a predominant role in the evolution of the southern capuchino radiation. Functional studies of specific variants in these divergent genomic regions will help clarify how novel allele associations could lead to different plumage phenotypes. Though the ultimate fate of the incipient *S. iberaensis* species remains uncertain, our findings illustrate how phenotypically differentiated lineages can form and rapidly become reproductively isolated from co-occurring, syntopic species (Rosenblum et al. 2012; Lamichhaney et al. 2018). Our results suggest that the reshuffling of standing genetic variation can generate novel phenotypes that are targeted by sexual selection. Assortative mating based on these traits may maintain species boundaries early in speciation while subsequent reproductive barriers accumulate.

4.5 MATERIALS AND METHODS

Field methods

We have carried out extensive field work in Iberá National Park (Argentina) since 2007, encountering individuals of all seven southern capuchino seedeaters that breed in the region. In the San Nicolás Reserve (28° 07' 41.4" S, 57° 26' 04.7" W), where this study took place, our group has conducted studies on the breeding ecology of capuchinos since 2014. During the study, *S. hypoxantha* and *S. iberaensis* were the only two southern capuchino species observed successfully breeding in San Nicolás. From November 2016 – January 2017 and October – December 2018, we located and monitored 128 nests of the two species (*S. hypoxantha*: $N = 65$; *S. iberaensis*: $N = 63$). We collected blood samples from the brachial veins of 126 adults and 77 nestlings (*S. hypoxantha*: $N = 40$ adult males, 23 adult females, 40 nestlings; *S. iberaensis*: 42 adult males, 21 adult females, 37 nestlings), as well as tissue samples from two unhatched eggs of *S. hypoxantha* and one unhatched egg of *S. iberaensis* (Table S4.1). In addition, we collected feather samples from four plumage patches across the body (crown, throat, belly, and rump) of individuals of *S. hypoxantha* ($N = 46$ males, 22 females) and *S. iberaensis* ($N = 41$ males, 20 females) to examine plumage coloration (described in the *Feather coloration* section). Males were attracted with playback and captured with mist nets during the nest construction, egg laying, and nestling provisioning stages, while females were captured at the nest during nestling provisioning. We measured and banded each individual with a numbered aluminum band and unique combination of colored leg bands prior to release. Blood samples were stored in lysis buffer and DNA was extracted with DNeasy blood and tissue kits (Qiagen, CA, USA) for all subsequent genomic analyses. From October – December 2019, we carried out an additional behavioral experiment in the San Nicolás Reserve (described in the *Behavioral experiment*

section) to test the importance of song and plumage coloration in species recognition and pre-mating isolation between *S. hypoxantha* and *S. iberaensis*.

Capuchino seedeaters are austral migrants that breed in the Iberá wetlands and migrate northwards during the non-breeding season (da Silva 1999). Over the course of the breeding season in 2018, we re-sighted nine individuals (16%) of *S. hypoxantha* ($N = 4$) and *S. iberaensis* ($N = 5$) of the 56 adults that were banded in 2016, two breeding seasons prior. In addition, in 2019, we re-sighted 22 banded males (26%) of *S. hypoxantha* ($N = 12$) and *S. iberaensis* ($N = 10$) of the 86 males banded from 2016-2018. Almost all of the re-sighted males were holding territories in the same geographic area of the study site as in previous breeding seasons. This relatively low recapture rate, but high philopatry, may be attributed to low inter-annual survival, and demonstrates high turnover in the individuals that are present at the breeding site across different years. The combination of a low recapture rate and the fact that females are indistinguishable to the human eye makes it difficult to quantify assortative mating by directly tracking the mating decisions of banded individuals across years (see *Assortative mating* section).

Whole-genome resequencing and variant discovery

We generated shotgun short-read whole-genome sequences for 37 individuals of *S. hypoxantha* ($N = 8$ males, 8 females) and *S. iberaensis* ($N = 12$ males, 9 females). Whole-genome resequencing generated over 860 million paired-end reads with a length of 151 bp, producing an expected per-individual coverage ranging between 3.9× and 10.7× (median: 5.4×; Table S4.2).

We evaluated the quality of individual libraries with FastQC (version 0.11.7) and used AdapterRemoval (version 2.1.7) to trim adapter sequences, filter by quality, and merge

overlapping paired-end reads (Andrews 2010; Schubert et al. 2016). The filtered data were aligned to a previously assembled reference genome of *S. hypoxantha* (Campagna et al. 2017) using the 'very-sensitive-local' option in Bowtie2 (version 2.3.4), and alignment statistics were subsequently obtained using Qualimap (version 2.2.1) (García-Alcalde et al. 2012; Langmead and Salzberg 2012). A high percentage of reads aligned to the reference genome (*S. hypoxantha*: mean = 98.6 ± 0.2% SD, *S. iberaensis*: mean = 98.6 ± 0.1% SD), and the average depth of coverage following filtering and alignment was 5.6× per sample (range: 3.8×-10.3×; Table S4.2).

We used SAMtools (version 1.7) to convert SAM to BAM files and sort and index the data (Heng et al. 2009). We then marked PCR duplicates with Picard Tools (version 2.17.10) and used HaplotypeCaller in GATK (version 3.8.0) to perform SNP variant discovery and genotyping (McKenna et al. 2010; Broad Institute 2018). The following hard filtering parameters were used to exclude variants in GATK: QD < 2.0, FS > 60.0, MQ < 20.0, and ReadPosRankSum < -8.0. We additionally used VCFtools (version 0.1.13) to filter out variants that had a minor allele frequency of less than 8% (retaining alleles present in at least three homozygous individuals), a mean depth of coverage lower than 2 or greater than 50, more than 20% missing data, or were not biallelic (Danecek et al. 2011). This pipeline produced 13,254,970 SNPs, and the average percent of missing data was 5% per individual.

Population genomics

To search for regions of elevated differentiation between *S. hypoxantha* and *S. iberaensis*, we computed average F_{ST} values for non-overlapping 5-kilobase (kb) windows, as well as individual SNPs, using VCFtools. Peaks of divergence were identified as 5-kb windows with elevated genomic divergence that contained at least one individual SNP with F_{ST} greater than 0.85. This

criterion focused on the strongest putative targets of divergent selection, though it may have excluded regions under selection that contained genes of small effect. We identified three peaks of divergence between the two species by building Manhattan plots and conducting principal component analyses (PCAs) of the genomic data with the packages *qqman* and *SNPRelate* in R version 3.5.2 (Zheng et al. 2012; R Core Team 2018; Turner 2018). In addition, we estimated D_{XY}, an absolute measure of divergence, over non-overlapping 5-kb windows for the three peaks of divergence using the custom script popgenWindows.py (https://github.com/simonhmartin/genomics_general). We explored patterns of linkage disequilibrium in these regions by calculating the r^2 statistic in plink version 1.9 (Purcell et al. 2007).

The reference genome was assembled to contigs from short-read shotgun and mate-pair libraries, and subsequently assembled to scaffolds using long-read data from Pacific Biosciences sequencing (Campagna et al. 2017). Although some of these scaffolds are large, they are not assembled to chromosome level. To assign scaffolds with divergence peaks to chromosomes, we aligned them to the zebra finch assembly (Taeniopygia_guttata-3.2.4) using the Satsuma synteny model from Satsuma version 3.1 (Grabherr et al. 2010). We assigned each scaffold to the chromosome with the top hit and examined the results with MizBee (Meyer et al. 2009). Finally, we referred to the annotated *S. hypoxantha* genome in (Campagna et al. 2017) to compile a list of genes within 50 kb of each divergence peak and searched for these annotations of interest in the UniProt (https://www.uniprot.org/) and Human Gene databases (https://www.genecards.org) to identify genes in the regions of elevated differentiation. The gene *OCA2* is adjacent to *HERC2* in the zebra finch, yet was not annotated in the *S. hypoxantha* reference genome. We located the *OCA2* coordinates in our reference genome by aligning the zebra finch mRNA

(XM_032749285.1) using BLAST (Altschul et al. 1990). To search for areas that could play an important role in regulating the expression of *OCA2* (*i.e.,* cis-regulatory elements), we assessed the level of conservation of the intergenic region between *OCA2* and *HERC2*, the area of the genome showing the highest differentiation between *S. iberaensis* and *S. hypoxantha*, with respect to more distantly related birds by using the Bird PhastCons track (Siepel et al. 2005) from the UCSC (University of California Santa Cruz) genome browser (Kent et al. 2002). Bird PhastCons scores are derived from a multigenome alignment of the budgerigar (*Melopsittacus undulatus*), zebra finch (*Taeniopygia guttata*), chicken (*Gallus gallus*), and turkey (*Meleagris gallopavo*) genomes, and represent the probability that a nucleotide belongs to a conserved element (ranging from 0 to 1). Areas that are highly conserved among distantly related species may contain regulatory elements that are important for controlling gene expression. We aligned the ~37 kb of sequence to the medium ground finch (*Geospiza fortis*) reference genome (geoFor1; to which the PhastCons scores were mapped) using BLAT (Kent 2002) with a 96.4% identity.

In addition, we assembled full mitochondrial genomes from the filtered whole-genome sequences belonging to *S. hypoxantha* and *S. iberaensis* individuals with MITObim 1.9.1 (Hahn et al. 2013), using the "quick" option and up to 40 iterations with the full mitochondrial genome from *Geospiza magnirostris* as a template (GenBank number NC_039770.1). We aligned the 37 individual sequences with an average length of 16,562 bp in Geneious version 10.1.3 (Kearse et al. 2012) and subsequently constructed an unrooted statistical parsimony network using PopART 1.7 (Leigh and Bryant 2015). In addition, we used the same methodology to generate a network based on the recovered COI DNA barcodes, which are frequently used for species identification (Hebert et al. 2003; Kerr et al. 2009; Campagna et al. 2010).

To generate phylogenetic hypotheses defined by the variants within the three divergence peaks in the context of the entire capuchino radiation, and compare these relationships to those in a whole-genome phylogeny of all capuchino species, we increased our genomic sampling to the ten southern capuchino species plus two outgroups. We combined the 37 whole-genome sequences obtained from *S. hypoxantha* and *S. iberaensis* with 72 additional individuals from nine capuchino species previously sequenced and published by Campagna *et al.* (Campagna et al. 2017), 12 new individuals from *S. ruficollis* sequenced on a lane of Illumina NextSeq 500 (paired-end, 151 bp), and six additional individuals sequenced on a lane of Illumina NextSeq 500 (mid-output mode, paired-end, 151 bp). All additional sequencing was performed at the Cornell University Biotechnology Resource Center (BRC). We assembled a VCF file (as described above in the *Whole-genome resequencing and variant discovery* section) with a total of 127 individuals (28 *S. hypoxantha*, 21 *S. iberaensis*, 15 *S. ruficollis*, 12 *S. pileata*, 12 *S. palustris*, 12 *S. melanogaster*, 12 *S. nigrorufa*, 4 *S. bouvreuil*, 4 *S. hypochroma*, 3 *S. cinnamomea*, and 2 individuals each of *S. minuta* and *S. castaneiventris* as outgroups). We applied the same hard filters as described above and subsequently retained variants that were present in 80% of all individuals, had a depth of coverage between 4 and 50, and a minor allele count of at least four. This combined dataset contained 32,993,511 SNPs after filtering. We explored the relationships among individuals and species by performing a PCA in *SNPRelate* and used VCFtools to create three additional files with subsets of SNPs from the regions defined by each divergence peak (1760 SNPs for the peak on scaffold 430, 1040 SNPs for the peak on scaffold 257, and 13 SNPs for the peak on scaffold 762). We used RAxML version 8.2.4 (Stamatakis 2014) to produce maximum likelihood phylogenies from the variants of each of the three divergence peaks, implementing the "ASC_GTRGAMMA" model and the Lewis correction for ascertainment bias.

For the more computationally demanding whole-genome phylogeny, we used RAxML-ng version 0.9.0 (Kozlov et al. 2019) and the "GTR+G+ASC_LEWIS" model. RAxML-ng used the 28.2 million SNPs that had the minor allele in homozygosity in at least one individual (from the total of ~33 million variants in the dataset). This analysis ran for approximately 1600 clock hours on all 64 cores of a computer with 512 Gb of RAM. Despite not converging on a single best phylogeny, an inspection of trees from the final search rounds showed very little variation, with only minor changes at the tips of the tree. We therefore generated a smaller dataset by applying more stringent filtering parameters (85% of individuals present at a locus and a minimum minor allele frequency of 10%), which retained 6,283,771 SNPs. We ran RAxML-ng on this dataset under the same conditions as described above, except that we used a parsimony tree as a starting point. This strategy converged on a single best tree, and both datasets (28.2 million and 6.3 million SNPs) produced comparable topologies.

Double-digest restriction site-associated DNA sequencing

To determine species identity for individuals without whole-genome data, assign paternity, and analyze patterns of assortative mating, we sequenced 206 individuals (126 adults, 77 nestlings, and 3 unhatched eggs from 23 nests of *S. hypoxantha* and 20 nests of *S. iberaensis*) in two separate sequencing runs following the double-digest restriction site-associated DNA (ddRAD) sequencing protocol detailed in (Thrasher et al. 2018).

We used FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/) to trim the 3' end of all reads to a length of 97 bp (FASTX Trimmer) and eliminated sequences (FASTX Quality Filter) if at least one base had a Phred score below 10 (90% call accuracy) or more than 5% of the bases had a score below 20 (99.9% call accuracy). We aligned the ddRAD data to the

reference genome of *S. hypoxantha* using the 'sensitive' option in Bowtie2 (version 2.3.5) and sorted and indexed the data with SAMtools (version 1.9). We then used the *gstacks* and *populations* modules of Stacks (version 2.3) to call variants and remove loci that were present in fewer than 80% of individuals (Catchen et al. 2013). The effective per-sample coverage was 31.8× ± 15.1× (mean ± SD). This pipeline produced a VCF file containing 61,484 SNPs across the 206 individuals.

Sanger sequencing of a region within the divergence peak on scaffold 430

To investigate the genomic architecture of phenotypic differences between the species in more detail, we developed a pair of primers (forward: 5'-ATTGCTGGTGTCTCCTTATTGA-3'; reverse: 5'-ATGTCCCTTTGGCTGTCTG-3') to sequence a ~700-bp region on scaffold 430 (11,028,673 to 11,029,376 bp) that contained 12 highly divergent SNPs ($F_{ST} > 0.85$) and an additional three SNPs with $F_{ST} > 0.79$. We amplified the divergent region via PCR for 165 individuals ($N = 87$ adults, 77 nestlings, and one unhatched egg) with GoTaq colorless master mix (Promega, WI, USA) and the following thermal cycle profile: 3 min at 95 °C, followed by 25 cycles of 30 sec at 95 °C, 30 sec at 61 °C, and one min at 72 °C, and finally 5 min at 72 °C. The PCR product was Sanger sequenced in both directions with the same primers used for amplification at the Cornell University BRC.

We used Unipro UGENE version 1.32.0 to trim primers and edit the Sanger sequences (Okonechnikov et al. 2012) and combined the information from these sequences with variants obtained through whole-genome sequencing to determine the genotypes of 202 individuals at 15 SNPs that showed high levels of differentiation within the peak on scaffold 430. We calculated F_{ST} values at each site using VCFtools and subsequently phased and imputed missing data

(~5.2% or 158 out of 3060 genotypes, with a mean probability of 0.996 ± 0036 SD) using BEAGLE version 3.3.2 (Browning and Browning 2007). This resulted in 404 haplotypes, two per individual. We explored the relationships between individuals at these sites by calculating a distance matrix in the R package *vegan* (Oksanen et al. 2010) and plotting it with the function *phylo.heatmap()* from the R package *phytools* (Revell 2012). We also compared these results to three similar plots derived from the *S. iberaensis* ($N = 21$) and *S. hypoxantha* ($N = 16$) individuals for which we had whole-genome sequencing data. We produced one plot for all SNPs found in the divergence peak on scaffold 430, one plot for the 64 SNPs with $F_{ST} > 0.79$ found in the same region, and a third plot for the 13 SNPs with $F_{ST} > 0.79$ found in the divergence peak on scaffold 257. We used an F_{ST} cutoff of 0.79, as the segment selected for PCR amplification included SNPs with this level of divergence.

Assortative mating

We analyzed patterns of social pairing from the whole-genome data ($N = 17$ social pairs) by first creating a tree of individuals using SVDquartets, implemented in PAUP*. SVDquartets is a coalescent-based method that compares possible quartet topologies for a set of four taxa, selecting the topology with the lowest score (Chifman and Kubatko 2014). In addition, we used the R package *qqman* to create Manhattan plots comparing the level of differentiation between species for males and females independently in each of the three divergence peaks. We evaluated whether individuals that formed a social pair grouped together on the tree, as expected if assortative mating contributes to reproductive isolation between *S. hypoxantha* and *S. iberaensis*, and showed elevated levels of differentiation in the same genomic regions.

For all social pairs ($N = 40$), including those with whole-genome data, we used the ddRAD pipeline to assign individuals to species and calculate the number of observed conspecific and heterospecific pairings. Specifically, we conducted a PCA of the genomic data using the *SNPRelate* package in R and evaluated if 1) individuals clustered by species in a PCA, and 2) socially or genetically determined male-female pairs (see *Paternity analysis* section) grouped together by species based on their diagnostic genomic PC1 scores.

Paternity analysis

We further filtered the VCF file from the ddRAD pipeline using the *populations* module of Stacks (version 2.3) to remove loci that had a minor allele frequency of less than 0.25, an observed heterozygosity greater than 0.7, or were present in fewer than 95% of individuals. We restricted the analysis to the first SNP per locus and used VCFtools to remove loci that were not in Hardy-Weinberg equilibrium or had a mean depth of coverage below 20. This pipeline produced a VCF file containing 281 highly informative loci across the 206 individuals that we used for paternity analysis.

Following filtering, we converted the VCF file to a format compatible with CERVUS 3.0.7, which takes a likelihood approach to assign paternity from SNP data (Kalinowski et al. 2007). CERVUS calculates the natural logarithm of the likelihood ratio (LOD score) for each potential pairing by comparing offspring genotypes to the genotypes of candidate parents and random individuals in the population. The LOD score thus estimates the relative likelihood that a sampled offspring was sired by a candidate father rather than a random male in the population. In addition, the program conducts a simulated parentage analysis using population allele

frequencies and the proportion of candidate parents sampled in the dataset to calculate the critical differences in LOD scores necessary to assign paternity with either 80% or 95% confidence.

To determine critical LOD scores, we simulated paternity assignments for 100,000 offspring (the recommended number) using the following parameters: 122 candidate males, 67% of candidate males sampled, and the default of 1% of loci mistyped. We approximated the proportion of candidate males sampled by estimating the number of males of both species that held neighboring territories to the sampled males but were never caught. The proportion of typed loci for the simulation was 0.972. As known mothers (confirmed by catching females at the nest) were sampled for 88% of offspring, we included known mothers in the analysis and evaluated CERVUS assignments using trio LOD scores, which take into account potential genotyping errors and the genotypes of known mothers when assigning paternity. Our total sample included 82 candidate males. We accepted assignments if the number of mismatches between the assigned male and his offspring was less than or equal to the maximum observed number of mismatches between a mother and her known offspring (as in Thrasher et al. (2018); max = 8, < 3% of 281 loci; Table S4.4). We assigned 51 of 77 nestlings (66%) to a candidate father with 95% confidence.

Feather coloration

We collected feathers from four plumage patches across the body (crown, throat, belly, and rump) from 68 individuals of *S. hypoxantha* (46 males and 22 females) and 61 individuals of *S. iberaensis* (41 males and 20 females) to examine plumage coloration. We stacked 10-15 feathers from each plumage patch on a non-reflective background surface (Flock Paper, Edmund Optics) to mimic their placement on the body of the bird. Reflectance data were generated relative to a

white standard (WS-1-SL, Ocean Optics) and a dark standard (all light omitted) with an Ocean Optics Flame spectrometer connected to an Ocean Optics PX-2 pulsed xenon light source. We used the OceanView software package (version 1.6.7, Ocean Optics) to record the reflectance data, averaging 20 scans per measurement. For each plumage patch, we took three measurements per individual and averaged the measurements prior to subsequent analysis. We used the R package *pavo* to compare the reflectance curves and degree of overlap in tetrahedral color space for each plumage patch between individuals of *S. hypoxantha* and *S. iberaensis* (Maia et al. 2013).

Behavioral experiment

From October – December 2019, we located males of *S. hypoxantha* ($N = 40$) and *S. iberaensis* ($N = 36$) that were actively singing on their territories in the San Nicolás Reserve and carried out a behavioral experiment in which we presented them with the following five treatments: (1) conspecific mount and song, (2) heterospecific capuchino mount and conspecific song, (3) conspecific mount and heterospecific capuchino song, (4) heterospecific capuchino mount and heterospecific capuchino song, and (5) heterospecific control mount and song. While *S. hypoxantha* and *S. iberaensis* form social pairs from October-November, our paternity data indicate that extra-pair mating continues throughout December in the San Nicolás Reserve. Male responses to mount presentation and song playback are often used to infer the importance of pre-mating isolation between divergent taxa (Irwin et al. 2001; Grant and Grant 2002a,b; Balakrishnan and Sorenson 2006; Uy et al. 2009b), as numerous studies have found that the traits used by males to recognize sexual competitors are also employed in female mate choice (Baker and Baker 1990; Baker 1991; Patten et al. 2004). For the heterospecific control, we followed the

methods of Benites et al. (2015) and used *S. collaris*, which is closely related to our focal species but not a capuchino seedeater (Mason and Burns 2013; Benites et al. 2015). *S. collaris* breeds in sympatry and occupies a very similar ecological niche to *S. hypoxantha* and *S. iberaensis* (Di Giacomo and Abril 2005; Di Giacomo and Krapovickas 2005). The heterospecific control treatment thus attempts to discriminate between aggressive responses to ecological and sexual competitors, given that all three species are grassland birds that feed on the seeds of tall grasses, such as *Paspalum durifolium* (Poaceae) and *Andropogon lateralis* (Poaceae), which dominate the landscape in the San Nicolás Reserve (Turbek et al. 2019; del Hoyo et al. 2020). In particular, an elevated response to conspecific traits relative to the stimuli of the heterospecific capuchino and control would indicate that 1) capuchinos recognize members of their own species as sexual competitors, and 2) the conspecific traits that elicit an elevated response are involved in male-male competition and potentially female choice (Baker and Baker 1990; Baker 1991; Patten et al. 2004). In contrast, a similarly aggressive response to conspecific and heterospecific capuchino stimuli would suggest that capuchinos do not discriminate between *S. hypoxantha* and *S. iberaensis*, recognizing males of both species as sexual and/or ecological competitors. Finally, an aggressive response to the control *S. collaris* stimuli would suggest that this more distantly related species, which is not a sexual competitor, elicits a response because it is recognized as an ecological competitor.

We recorded the geographic coordinates of each trial and tested focal males with as many treatments as possible (up to five treatments) by returning to the same geographic location multiple times. Trials performed with the same focal males were separated by at least one day, and the order in which treatments were presented was randomized. In addition, we randomized the order in which stimuli were presented across trials and ensured that the mounts (two of *S.*

hypoxantha, two of *S. iberaensis*, and two of *S. collaris*) and playback files (10 of *S. hypoxantha*, 10 of *S. iberaensis*, and five of *S. collaris*) were presented an equal number of times. Sixteen of the 76 focal males (21%) were color-banded from our field work in previous years. In addition, capuchino seedeaters exhibit a considerable degree of intraspecific variation in plumage coloration, which is likely associated with age, and only sing within their territories. Thus, by returning to the same location where a male was previously observed singing, and using color bands or plumage to identify individuals, we could be confident that the same individual was tested in subsequent trials.

During each trial, we located the focal male and set up the mount approximately 1-2 m off the ground < 35m from the focal male on a thin pole near vegetation suitable for perching. We hid a compact speaker (JBL Flip 5) in the vegetation under the mount and connected the speaker to a phone through Bluetooth to start the playback recordings. Each trial lasted a total of five minutes (the duration of the playback file), with the same observer (always SPT, for consistency in scoring behavioral responses) standing 20 meters away from the mount. We generated video recordings of each trial with a DSLR camera (Canon EOS 7D) and dictated vocalizations and behaviors into the camera during the trials. We recorded the following behavioral responses: the number of flights and amount of time spent at various distances from the mount, the number of attacks and amount of time spent attacking the mount, and the amount of time spent singing by each focal male, using the 2m pole on which the mount was placed to estimate distance from the mount. While females do not assist with territorial defense in capuchino seedeaters, we noted whether or not a female was observed during each trial in case female presence influences male response to territorial intrusion. Females were observed in 53 trials (22%). In total, we presented 32 males (16 *S. hypoxantha* and 16 *S. iberaensis*) with all five

treatments and 44 males (24 *S. hypoxantha* and 20 *S. iberaensis*) with fewer than five treatments, for a total of 240 trials (24 trials per treatment for each species).

We ran a PCA on the correlation matrix of the behavioral response variables using the R package *vegan* to reduce the dimensionality of the behavioral data. The PCA identified three axes of variation (eigenvalues > 1) that collectively explained 79% of the variation in behavioral responses (PC1: 43%, PC2: 20%, PC3: 16%). All input variables associated with male aggression (*e.g.,* number of flights near the mount, proportion of time spent near the mount, and number of attacks at the mount) loaded positively on PC1, while proportion of time spent singing and proportion of time spent over six meters from the mount loaded negatively on PC1 (Figure S4.15C), indicating that PC1 represented a reasonable overall summary of aggression. We therefore extracted PC1 to generate a response intensity score for each trial. In addition, we classified all trials as displaying attack behavior ('1') or not ('0'), with attack behavior defined as either swooping at or making direct contact with the mount, to examine a direct indicator of aggression. We carried out parallel analyses with response intensity (*i.e.,* PC1) and attack behavior as dependent variables using R version 3.5.2 (R Core Team 2018) and fit generalized linear mixed models (GLMMs) with the R packages *lme4* (for linear mixed models) and *glmmTMB* (for mixed logistic regression models) to analyze the responses of territory owners to mount presentation and song playback (Bates et al. 2015; Brooks et al. 2017).

Because territorial males typically did not respond to the heterospecific control (*S. collaris*) stimuli, we ran separate analyses with and without the control trials (as in Uy et al. (2009a)). Excluding the control trials, we first ran GLMMs examining the effects of species, plumage, and song (fixed effects) on 1) response intensity (PC1) and 2) attack behavior (whether or not the mount was attacked at any point during the trial). We used a mixed logistic regression

model with a binomial distribution and logit link function to model attack behavior, which had a binary outcome (0 or 1). Preliminary models included treatment order, male ID, female presence (0 or 1), mount ID, and playback ID as random effects, with mount ID and playback ID nested within each plumage and song type, respectively, in order to control for repeated measures from individuals and mount/playback exemplar effects. We calculated a 95% confidence interval around the estimated standard deviation explained by the random effects using the *confint()* function from the R stats package. We excluded random effects from the model if the lower end of the confidence interval reached zero, indicating that the effect did not account for variation in the model (*e.g.,* treatment order, mount ID, and playback ID), retaining male ID in the model of attack behavior and male ID and female presence in the model of response intensity.

In addition, we included the heterospecific control trials to run a GLMM for each focal species that tested the effect of treatment group on response intensity (PC1), incorporating male ID and female presence as random effects. We used the R package *emmeans* to run post-hoc pairwise comparisons between treatment groups using Tukey's honestly significant difference (HSD) test (Lenth et al. 2018). Again, treatment order, mount ID, playback ID, and female presence (in the case of *S. hypoxantha*) had 95% confidence intervals that reached zero when included as random effects in the preliminary models and were therefore excluded from the final analyses. We verified the assumptions of the linear mixed models by generating Q-Q plots and plotting the residuals versus the fitted values.

Abundance estimates

To estimate whether *S. iberaensis* has increased in abundance across its breeding range since the species' first records in 2001 (Di Giacomo and Kopuchian 2016), we downloaded eBird data

through January 2020 from the February 2020 release of the eBird Basic Dataset (eBird Basic Dataset 2020). The downloaded dataset contained all available data for the control used in the behavioral experiment (*S. collaris*) and the seven capuchino species (*S. iberaensis*, *S. hypoxantha*, *S. cinnamomea*, *S. palustris*, *S. pileata*, *S. ruficollis*, and *S. hypochroma*) that breed in Iberá National Park (BirdLife International 2020), as well as the Sampling Event Data (needed to infer non-detection records). eBird is an online database where scientists, researchers, and amateur naturalists can upload avian observations (Sullivan et al. 2014). We filtered the data for each species using the R package *auk* (Strimas-Mackey et al. 2017) to exclude records from incomplete checklists (*i.e.,* checklists in which some identified species were not reported) and observations that fell outside of a bounding box around the area encompassing all observations of *S. iberaensis* in Argentina and Paraguay (the central range of *S. iberaensis*), retaining only a single checklist from each set of non-independent ("shared") checklists. We then inferred non-detection records (*i.e.,* "zero-filled the data") using the *auk* package to create presence/absence data for each species. To more precisely define the spatial area of interest, we converted the presence-only data from *S. iberaensis* into a spatial object and generated a convex hull polygon around the distribution of *S. iberaensis* records using the function *gbuffer()* in the R package *rgeos* (Bivand and Rundel 2017), adding a buffer of 1 map unit (with data in a South America Albers Equal Area Conic projection) outside of these locations; we only retained records of observations that fell within this polygon of interest. We then placed temporal restrictions on the remaining data, only retaining records from October-February, when capuchino seedeaters are present on the breeding grounds, and records beginning with the austral summer that spanned the years 2013-2014, the first summer for which multiple observations of *S. iberaensis* existed in the eBird database.

After processing the data, we examined whether there have been any systematic changes in the reporting rates of the seven capuchino seedeater species over the past decade in order to assess whether *S. iberaensis* has increased in prevalence relative to other capuchino species. We modeled changes in prevalence by fitting generalized additive models (GAMs; using R package *mgcv* (Wood 2011)), in which the probability of reporting of the focal species was modeled as a function of the calendar year at the end of each austral summer. GAMs are able to identify arbitrary, continuous patterns of change through time, rather than forcing specific patterns onto the data. We used additional smoothing terms to account for variation in observation effort as described by the following variables: observation date, distance traveled, and duration of the observation period. We compared the GAM results for each capuchino species to determine whether 1) there was a significant change in the reporting probability of each species over time and 2) whether reporting probability consistently increased from 2014-2019.

CHAPTER 5

VARIABLE SIGNATURES OF SELECTION DESPITE CONSERVED RECOMBINATION LANDSCAPES EARLY IN SPECIATION

5.1 ABSTRACT

Recently diverged taxa often exhibit a pattern known as heterogeneous genomic divergence, characterized by regions of elevated differentiation in an otherwise homogeneous background. While divergence peaks are generally interpreted as regions responsible for reproductive isolation, they can also arise due to the individual or combined effects of background selection, selective sweeps in allopatry, and variation in recombination and mutation rates across the genome. To investigate the association between patterns of recombination and landscapes of genomic differentiation during the early stages of the speciation process, we generated fine-scale recombination maps for six southern capuchino seedeaters (*Sporophila*) and two White Wagtail subspecies (*Motacilla alba*), two recent avian radiations in which divergent selection on pigmentation genes has likely generated peaks of differentiation. We compared the recombination maps of these recent radiations to those of the Collared (*Ficedula albicollis*) and Pied Flycatcher (*Ficedula hypoleuca*), non-sister taxa characterized by moderate genomic divergence and a heterogenous genomic landscape that has been shaped in large part by background selection. Although we found that recombination landscapes were conserved within all three systems, we documented a weaker correlation between recombination rate and genomic differentiation in the recent radiations than in the more divergent Collared and Pied Flycatchers. All divergence peaks between capuchinos and wagtails were located in regions with lower than average recombination rates, and a large portion of the divergence peaks in capuchinos fell in

regions of exceptionally reduced recombination. Thus, the association among genes in these regions may have been protected early in divergence, facilitating rapid diversification. Nonetheless, divergence peaks are present in unique combinations when comparing different capuchino species, despite largely conserved recombination landscapes, suggesting that regions of elevated differentiation have not been generated by variation in recombination rate alone.

5.2 INTRODUCTION

During population divergence and speciation, the level of genetic differentiation among closely related taxa can vary widely across the genome, a pattern referred to as heterogeneous genomic divergence (Nosil et al. 2009). Regions of elevated differentiation are often thought to harbor barrier loci responsible for reproductive isolation, while regions of low differentiation are attributed to the homogenizing effects of gene flow or a lack of differentiation in allopatry (Turner et al. 2005). Accordingly, researchers routinely rely on genomic scans (*e.g.,* F_{ST} outlier tests) to identify loci whose genetic differentiation exceeds neutral expectations and that are therefore assumed to be under divergent selection (Nosil et al. 2008). However, a growing body of literature indicates that genomic landscapes of differentiation are much more complex and difficult to interpret than previously thought (Cruickshank and Hahn 2014; Burri et al. 2015). For example, other selective forces like background selection and selective sweeps due to local adaptation in alloaptry, as well as variation in recombination and mutation rates, can contribute to patterns of heterogeneous genomic divergence by altering levels of within-population diversity, thereby confounding the findings of genomic scans and potentially yielding high false positive rates when searching for speciation loci under divergent selection (Charlesworth et al. 1993; Ravinet et al. 2017; Booker et al. 2020).

While chromosomal rearrangements and their corresponding suppression of recombination in heterozygotes have long been known to promote genomic differentiation (Rieseberg 2001), the contribution of genome-wide variation in recombination rate to patterns of heterogeneous genomic divergence during speciation has received comparatively less attention (Butlin 2005; Roesti et al. 2012). Recently, the greater availability of whole-genome sequence data and advances in algorithmic methods for estimating recombination have provided the opportunity to investigate the chromosomal features that generate variation in patterns of recombination across the genome (Smukowski and Noor 2011). Studies estimating recombination rate through cytological, linkage mapping, and linkage disequilibrium-based methods have documented substantial variation in recombination rate among different genomic regions, according to factors such as chromosome type (*e.g.,* autosome vs. sex chromosome), distance to centromere and chromosome center, gene density, and GC content, and have demonstrated that recombination estimates can differ among closely related taxa (Smukowski and Noor 2011; Berner and Roesti 2017; Stapley et al. 2017). In addition, they have revealed the presence of recombination hotspots, or regions that shuffle genetic variation at higher rates than the rest of the genome (Lichten and Goldman 1995; Singhal et al. 2015), emphasizing the need to consider variation in local recombination rate when interpreting patterns of elevated genomic differentiation.

In genomic areas of low recombination, the reduction in diversity at linked sites due to background selection and selective sweeps (*i.e.,* linked selection) and the effects of drift can both extend over physically larger regions than in areas of high recombination, elevating estimates of genomic differentiation between populations. Indeed, empirical studies in a variety of systems have detected a negative relationship between F_{ST} and recombination rate (Stephan et al. 1998;

Keinan and Reich 2010; Renaut et al. 2013), supporting the idea that linked selection in regions of low recombination can generate heterogeneous genomic landscapes even in the absence of selection against gene flow. However, few studies have investigated the association between patterns of recombination and landscapes of genomic differentiation during the earliest stages of the speciation process (Smukowski and Noor 2011). Often, incipient species are only differentiated by a handful of genomic regions (Martin et al. 2013; Marques et al. 2016). In several avian systems, in particular, the few regions of elevated genomic differentiation between closely related taxa contain genes involved in plumage patterning and coloration, suggesting that divergent selection on plumage traits used in mate choice and territorial defense has played a key role in speciation (Poelstra et al. 2014; Toews et al. 2016; Campagna et al. 2017; Cooper and Uy 2017; Stryjewski and Sorenson 2017; Semenov et al. 2021; Turbek et al. 2021). Nonetheless, in systems such as these, it remains unclear whether regions containing pigmentation genes are also associated with regions of particularly high or low recombination.

To further investigate the interplay between linked selection and divergent selection at various stages of the speciation process, we generated fine-scale recombination maps for southern capuchino seedeaters (*Sporophila*) and White Wagtails (*Motacilla alba*), two recent avian radiations with well-characterized genomic landscapes and a history of ongoing or recent gene flow. We compared the patterns observed in these radiations to those detected in the Collared (*Ficedula albicollis*) and Pied Flycatcher (*Ficedula hypoleuca*), non-sister taxa that exhibit moderate genomic divergence but still occasionally inbreed (Burri et al. 2015). Collared and Pied Flycatchers differ in male plumage coloration and song, as well as various life-history traits, and hybridize at low frequencies in areas of sympatry in eastern Central Europe and the Baltic Islands of Gotland and Öland (Qvarnström et al. 2010; Ellegren et al. 2012). While female

F1 hybrids are entirely sterile, male hybrids experience reduced fitness as a result of sexual selection against intermediate plumage patterns (Svedin et al. 2008). Thus, numerous traits contribute to pre- and post-zygotic isolation between the flycatcher species (Qvarnström et al. 2010).

In contrast to *Ficedula* flycatchers, southern capuchino seedeaters constitute one of the most recent and rapid avian radiations and comprise ten highly sympatric species that are ecologically and genomically similar, yet exhibit striking differences in male plumage patterning and song (Campagna et al. 2012, 2017). Many of the few genomic regions that differentiate the species contain pigmentation genes involved in the melanogenesis pathway (Campagna et al. 2017) and are associated with recent, species-specific selective sweeps (Hejase et al. 2020). In addition, two syntopic capuchino species (*S. hypoxantha* and *S. iberaensis*) mate assortatively in Iberá National Park in Corrientes, Argentina, where pre-mating isolation appears to be based on plumage coloration and song (Turbek et al. 2021). Collectively, these results suggest that capuchino seedeaters diversified through sexual selection on melanin-based plumage traits (Campagna et al. 2017). Similar to capuchino seedeaters, White Wagtails (*Motacilla alba*) are a recent radiation consisting of nine subspecies characterized by variation in plumage and body size, but little genomic divergence (Harris et al. 2018; Semenov et al. 2018). Two subspecies, *M. a. alba* and *M. a. personata*, form a hybrid zone in central Siberia, where social pairs mate assortatively by head plumage (Semenov et al. 2017), a trait with a simple genetic basis that maps to two small autosomal regions of elevated genomic differentiation (Semenov et al. 2021). Thus, divergent mate preferences, combined with selection on genes encoding plumage pigmentation, appear to underlie the maintenance of phenotypic differentiation in both capuchino seedeaters and White Wagtails.

The genomic landscapes of differentiation between *Ficedula* flycatchers, capuchinos, and wagtails are highly heterogeneous, with regions of elevated and reduced divergence (Ellegren et al. 2012; Campagna et al. 2017; Semenov et al. 2021; Turbek et al. 2021). However, while capuchinos and wagtails are characterized by an extremely low level of background genomic differentiation [capuchinos: genome-wide F_{ST} = 0.008 (Campagna et al. 2017), wagtails: genome-wide F_{ST} = 0.046 (Semenov et al. 2021)] punctuated by few divergence peaks, Collared and Pied Flycatchers exhibit a much higher background of differentiation (autosomal F_{ST} = 0.350) and multiple divergence peaks that are scattered throughout the genome (Ellegren et al. 2012). Most of these F_{ST} peaks have evolved repeatedly in the same genomic regions among independent lineages of *Ficedula* flycatchers, suggesting that some of the more prominent features of the heterogeneous genomic landscape between Collared and Pied Flycatchers evolved as a result of background selection and selective sweeps due to local adaptation in regions of low recombination rather than divergent selection on genomic regions involved in reproductive isolation (Burri et al. 2015). By comparing *Ficedula* flycatchers to the recent capuchino and wagtail radiations, we can therefore examine the degree of conservation of recombination rates among closely related species and assess the location of known divergence peaks relative to the recombination landscape at various stages of the speciation process. Background selection is thought to explain accentuated genomic differentiation among more divergent taxa, while positive selection may play a more prominent role in generating regions of elevated differentiation early in divergence (Burri 2017); however, few studies have evaluated this hypothesis by comparing landscapes of heterogeneous genomic differentiation along the speciation continuum.

Given that recombination breaks up favorable allele combinations, potentially imposing fitness costs early in divergence, we expected to observe an association between regions of low recombination and high F_{ST} in all three systems. While this pattern is well-established later along the speciation continuum due to the accumulating effects of background selection (Cruickshank and Hahn 2014; Burri 2017), and should therefore be supported in *Ficedula* flycatchers, it may not be detectable in the earliest stages of the speciation process. If the time to divergence in capuchino seedeaters and White Wagtails is too recent for this pattern to be detected, we expect little to no connection between recombination rate and genomic differentiation in capuchinos and wagtails, in contrast to *Ficedula* flycatchers. The genetic architecture of phenotypic differentiation in both capuchinos and wagtails is relatively simple, involving only a few narrow genomic regions, and could potentially persist regardless of broader (*e.g.,* chromosome-scale) recombination patterns. Finally, F_{ST} outliers in capuchino seedeaters and White Wagtails could fall in regions of high recombination. While recombination breaks apart co-adapted genetic combinations, it also generates new variability upon which natural selection can act (Ortiz-Barrientos et al. 2016) and reduces the linkage disequilibrium between loci under selection (*i.e.,* Hill–Robertson interference), thereby facilitating adaptation in natural populations (Hill and Robertson 1966; Comeron et al. 2008). High recombination in regions containing pigmentation genes could thus generate novel associations between genes and/or between elements in their regulatory network, leading to new plumage patterns, and make it easier for natural selection to operate on individual loci that contribute to phenotypic diversity.

5.3 MATERIALS AND METHODS

Sampling

To calculate recombination rates for each system, we used published whole-genome sequencing data from southern capuchino seedeaters (Campagna et al. 2017; Turbek et al. 2021), White Wagtails (Semenov et al. 2021), and *Ficedula* flycatchers (Burri et al. 2015). The capuchino dataset included 123 individuals of 10 southern capuchino seedeater species: the Tawny-bellied Seedeater (*S. hypoxantha*; n = 28), the Iberá Seedeater (*S. iberaensis*; n = 21), the Black-bellied Seedeater (*S. melanogaster*; n = 12), the Black-and-tawny Seedeater (*S. nigrorufa*; n = 12), the Marsh Seedeater (*S. palustris*; n = 12), the Pearly-bellied Seedeater (*S. pileata*; n = 12), the Chestnut Seedeater (*S. cinnamomea*; n = 3), the Rufous-rumped Seedeater (*S. hypochroma*; n = 4), the Dark-throated Seedeater (*S. ruficollis*; n = 15), and the Copper Seedeater (*S. bouvreuil*; n = 4), as well as four individuals of two outgroup species: the Chestnut-bellied Seedeater (*S. castaneiventris*; n = 2) and the Ruddy-breasted Seedeater (*S. minuta*; n = 2; Table S5.1). The Tawny-bellied Seedeater and Iberá Seedeater datasets included a combination of males and females, while only males were included for the other capuchino species. For White Wagtails, we examined sequences from ten males of *M. a. alba* and ten males of *M. a. personata* that were sampled from allopatric populations in western Siberia and Uzbekistan, respectively, as well as five individuals of the Forest Wagtail (*Dendronanthus indicus*) and four individuals of the Gray Wagtail (*Motacilla cinerea*) as outgroups (Table S5.1). The *Ficedula* dataset contained ten males each of the Collared (*F. albicollis*) and Pied Flycatcher (*F. hypoleuca*) that were sampled in a zone of sympatry in the Czech Republic, along with single individuals of two outgroup species, the Red-breasted Flycatcher (*F. parva*) and the Snowy-browed Flycatcher (*F. hyperythra*) (Table S5.1).

Whole-genome sequencing and variant calling

Raw sequencing reads were processed as described in Turbek et al. (2021) for capuchinos and *Ficedula* flycatchers and Semenov et al. (2021) for wagtails. Briefly, we used AdapterRemoval (capuchinos: v.2.1.7, flycatchers: v.2.3.1) to trim capuchino and *Ficedula* reads with a Phred quality score below 10, remove adapter sequences, and merge overlapping paired-end reads (Schubert et al. 2016). We aligned the capuchino reads to a reference genome of *S. hypoxantha* (Campagna et al. 2017) and the *Ficedula* reads to the chromosome-level collared flycatcher genome assembly FicAlb1.5 (Kawakami et al. 2014) using the very sensitive, local option in Bowtie2 (capuchinos: v.2.3.4, flycatchers: v.2.4.2) (Langmead and Salzberg 2012). For wagtails, we used the default settings in Trimmomatic v.0.39 to trim poor-quality reads and remove Illumina adapters (Bolger et al. 2014). We then used a chromosome-level assembly of *M. alba alba* (NCBI BioProject PRJNA662706) to align the reads using bwa mem (v.0.7.17-r1188) within the Sentieon wrapper (Freed et al. 2017).

Following alignment, we used SAMtools (v.1.7) to sort and index the capuchino and flycatcher data and Picard Tools (capuchinos: v.2.17.10, flycatchers: v.2.23.8) to mark PCR duplicates, realign around indels, and fix mate pairs (Heng et al. 2009; Broad Institute 2018). Variant calling was performed with HaplotypeCaller in GATK (capuchinos: v.3.8.0, flycatchers: v.4.1.9) (McKenna et al. 2010). We merged the resulting g.vcf files with CombineGVCFs and genotyped the data with GenotypeGVCFs. For wagtails, we ran duplicate removal, HaplotypeCaller (GATK 4.1), and GenotypeGVCFs (GATK 4.1) using the Sentieon wrapper tools (Freed et al. 2017). Next, we used VariantFiltration in GATK to filter out variants based on the following parameters: QD < 2.0, FS > 60.0, MQ < 30.0, MQRankSum < -12.5 and ReadPosRankSum < -8.0 for capuchinos and flycatchers and QD < 2.0, FS > 60.0, MQ < 40.0,

and MQRankSum < -12.5 for wagtails. Finally, we used vcftools (capuchinos: v.0.1.13, wagtails and flycatchers: v.0.1.15) to remove individuals with more than 20% missing data (3 *S. hypoxantha*, 3 *S. nigrorufa*, 2 *S. palustris*, 2 *S. pileata*) and exclude variants that had a minor allele frequency less than 0.05, more than 10% missing data, or were not biallelic (Danecek et al. 2011). We additionally filtered out sites that had a quality score below 20 and mean depth of coverage less than 4 or greater than 75 (for wagtails) and sites with a mean depth of coverage less than 2 or greater than 50 (for capuchinos and flycatchers). This pipeline yielded 10,638,161 SNPs after filtering for capuchinos, 37,215,097 SNPs for wagtails, and 15,072,062 SNPs for *Ficedula* flycatchers.

We filtered the capuchino dataset to remove species that lacked the necessary number of individuals to calculate recombination rate (*S. cinnamomea*, *S. hypochroma*, *S. ruficollis*, and *S. bouvreuil*) and created separate vcf files for each species in order to estimate recombination rate (for the ingroups) or infer the ancestral allele at each variable site (for the outgroups). Although the wagtail and *Ficedula* references genomes were assembled to chromosome level, our reference for *S. hypoxantha* was not a chromosome-level assembly. We therefore estimated recombination rate for each capuchino scaffold individually, reoriented and ordered the scaffolds according to their alignment to the zebra finch assembly (Taeniopygia_guttata-3.2.4), removing scaffolds that had a cumulative matching sum < 5 kb, and merged the LDhelmet output for scaffolds assigned to each chromosome.

Recombination rate estimation

We computed recombination rate over 50-kilobase (kb) windows using LDhelmet v.1.10 (Chan et al. 2012). LDhelmet infers recombination rates by examining patterns of linkage

disequilibrium (LD) across phased haplotypes and requires 1) an estimate of θ, the population mutation rate, 2) a set of phased haplotypes, 3) knowledge of the ancestral allele at each segregating site, and 4) a 4x4 mutation matrix containing counts of mutation types (Chan et al. 2012). We estimated the population mutation rate for each system by calculating Watterson's θ per-chromosome using a custom python script modified from Shanfelter et al. (2019). The script used the R package *PopGenome* (Pfeifer et al. 2014) to estimate Watterson's θ over 2-kb windows with a sliding window of 1 kb and averaged all windows together for each chromosome. Given that estimates of Watterson's θ were very similar among taxa within each system (± 0.001), we used the same population mutation rate estimate within capuchinos, wagtails, and *Ficedula* flycatchers. A previous study found that small changes in θ have little influence on the resulting recombination estimates generated in LDhelmet (Stukenbrock and Dutheil 2018). To infer haplotypes, we phased each chromosome (or scaffold for capuchinos) independently with read-aware phasing in ShapeIt2, which uses reads that span at least two heterozygous sites (*i.e.,* phase informative reads) to improve phasing accuracy (Delaneau et al. 2013). ShapeIt2 requires a minimum of ten individuals for phasing. Given that three individuals of *S. nigrorufa* had more than 20% missing data and were excluded from the vcf, we included one individual of *hypoxantha* with those of *S. nigrorufa* for the phasing step, as capuchino seedeaters exhibit little neutral genomic structure, and subsequently removed this *S. hypoxantha* individual from the phased vcf for downstream analyses. For capuchino species with more than ten individuals following filtering (*S. hypoxantha*, *S. iberaensis* and *S. melanogaster*), we excluded birds with more than 10% missing data prior to phasing each scaffold, as recommended to improve phasing accuracy. All wagtail and flycatcher individuals had low rates of missing data, and were therefore retained in the phasing analysis.

We used two outgroup species per system to infer the ancestral allele at each variable site. If the outgroup taxa were homozygous for the same allele at a particular locus, the nucleotide carried by the outgroups was assigned a prior probability of 0.91, while the remaining three nucleotides were assigned prior probabilities of 0.03 to account for uncertainty in ancestral allele reconstruction. For all other sites (for which a polymorphism was segregating among the outgroup species), each nucleotide was assigned a prior probability derived from the overall frequency of that particular nucleotide on the chromosome, as in Singhal et al. (2015). We estimated mutation matrices for each species by quantifying the total number of each type of mutation away from the ancestral allele at all positions for which an ancestral allele could be inferred. Ancestral allele probabilities and mutation matrices were generated using a custom perl script modified from Shanfelter et al. (2019).

We converted the vcf file for each ingroup species into SNP sequence and SNP position files with the flag '-ldhelmet' in vcftools. For each chromosome of interest, we used the *find_confs* module in LDhelmet with a window size of 50 SNPs (-w 50) to create one haplotype configuration file. Next, we generated one likelihood lookup table per chromosome with the *table_gen* command using the recommended grid of ρ values (-r 0.0 0.1 10.0 1.0 100.0) and our estimate of Watterson's θ for each system (capuchinos: 0.01, wagtails: 0.0045, flycatchers: 0.004). Additionally, we ran the *pade* module to generate Padé coefficient tables with our estimate of Watterson's θ (-t) and the recommended 11 coefficients (-x 11).

For the *rjmcmc* step, we used the default burn-in of 100,000 iterations and ran the Markov chain for 1,000,000 iterations to estimate recombination rate for each chromosome (or scaffold in capuchinos) individually over the recommended window size of 50 SNPs (-w 50 -burn_in 100000 -n 1000000). The *rjmcmc* procedure also requires specification of the block

penalty, a penalty assigned to the likelihood when the recombination rate changes across the genome. A previous study that used LDhelmet to generate recombination maps in birds found that a block penalty of 5 maximizes the power to detect hotspots, while a block penalty of 100 generates the most accurate recombination maps over longer distances (Singhal et al. 2015). As we were interested in broader-scale patterns of recombination, we ran each chromosome with a block penalty of 100 (-b 100) in all three systems. Finally, we extracted population-scaled recombination rates using the *post_to_text* command in LDhelmet. Recombination rates are reported in ρ/bp, where ρ is the population-scaled recombination rate ($4N_e r$) and r refers to the per-generation recombination rate (*i.e.,* the probability of a recombination event occurring during meiosis).

We estimated recombination rates for chromosomes containing the main divergence peaks between the focal species (1, 1A, 11, 20, and Z in capuchinos and 1A and 20 in wagtails). For *Ficedula* flycatchers, we generated recombination maps for all five of the analyzed chromosomes, as each chromosome contains a divergence peak between Pied and Collared Flycatchers (Ellegren et al. 2012).

Statistical analyses

For each chromosome, we filtered out recombination estimates between SNPs that were located greater than 5 kb apart, as SNP density was too low to trust these estimates, and averaged recombination rate (ρ) over 100-kb windows with a 100-kb step using the R package *WindowScanR* (Tavares 2020). In addition, we used vcftools to calculate F_{ST} over 25-kb windows between the species pairs within each system. To examine the extent to which recombination landscapes are conserved within systems, we calculated Pearson's correlation

coefficient (*r*) between the windowed ρ estimates of species pairs. We also computed Pearson's *r* between recombination rate and F_{ST} for each chromosome of interest to compare the association between regions of elevated genomic differentiation and patterns of recombination in *Ficedula* flycatchers and the recent capuchino and wagtail radiations. We note that the *p*-values associated with these correlations are affected by pseudoreplication, as loci are genetically linked and therefore not independent; however, we have decided to include these correlations in the study as a way to illustrate and summarize the genome-wide pattern. To determine whether divergence peaks in capuchino seedeaters and White Wagtails that are known to be under divergent selection fall in regions of exceptionally high or low recombination, we carried out randomization tests for *M. a. alba* vs. *M. a. personata* and the capuchino species that exhibited the highest (*S. nigrorufa* vs. *S. melanogaster*; mean genome-wide $F_{ST} = 0.008$) and lowest (*S. hypoxantha* vs. *S. iberaensis*; mean genome-wide $F_{ST} = 0.005$) genomic differentiation. For each chromosome, we calculated mean ρ within the divergence peaks and randomly chose regions of the same size (*i.e.,* with the same number of windows as the divergence peaks) 10,000 times, with replacement. We computed mean ρ for each randomly selected region and calculated the two-tailed *p*-value as 2 * the number of times the mean ρ of the randomly selected region fell below the mean observed ρ within the divergence peaks, divided by the number of replicates. All analyses were carried out in R v.3.6.2 (R Core Team 2018). The randomization tests examined the null hypothesis that the average ρ of the divergence peak was equal to the chromosome-wide average ρ. We did not include *Ficedula* flycatchers in the randomization tests given that the Collared and Pied Flycatcher have a much higher background level of genomic differentiation and, unlike capuchino seedeaters and White Wagtails, lack clear F_{ST} outliers.

5.4 RESULTS

Closely related species are characterized by similar recombination landscapes

Recombination landscapes were highly correlated within capuchino seedeaters, wagtails, and *Ficedula* flycatchers, indicating that recombination rates are fairly conserved among closely related avian species (Table 5.1, Figure S5.1). Correlation coefficients for wagtails, *Ficedula* flycatchers, and the capuchino species exhibiting the highest and lowest genomic differentiation are provided in Table 5.1, while correlation coefficients among all capuchino species are provided in Figure S5.2. Pearson's correlation across chromosomes ranged from 0.71-0.97 in capuchino seedeaters (5 chromosomes), 0.71-0.86 in White Wagtails (2 chromosomes), and 0.64-0.87 in *Ficedula* flycatchers (5 chromosomes). Although divergence peaks are present in unique combinations when comparing different capuchino seedeaters (see example in Figure S5.3) (Campagna et al. 2017), recombination rate was remarkably similar among capuchino species (Figure S5.1, Figure S5.2). Males and females can differ in their fine-scale recombination patterns (Sardell and Kirkpatrick 2020); however, *S. hypoxantha* and *S. iberaensis* exhibited similar recombination estimates as the other capuchino species despite the fact that they included a combination of males and females (Figure S5.1).

Table 5.1. Pearson's correlation (*r*) between the recombination landscapes of closely related species or subspecies of capuchino seedeaters, White Wagtails, and *Ficedula* flycatchers. Each row refers to a chromosome that contained at least one divergence peak when the two species were compared. $p < 0.05$ for all comparisons.

Comparison	**Chr**	**Pearson's *r***
S. hypoxantha vs. *S. iberaensis*	1	0.798
S. hypoxantha vs. *S. iberaensis*	11	0.914
S. hypoxantha vs. *S. iberaensis*	Z	0.798
S. nigrorufa vs. *S. melanogaster*	1	0.816
S. nigrorufa vs. *S. melanogaster*	1A	0.913
S. nigrorufa vs. *S. melanogaster*	11	0.882
S. nigrorufa vs. *S. melanogaster*	20	0.892
S. nigrorufa vs. *S. melanogaster*	Z	0.789

M. a. alba vs. *M. a. personata*	1A	0.864
M. a. alba vs. *M. a. personata*	20	0.709
F. albicollis vs. *F. hypoleuca*	1	0.828
F. albicollis vs. *F. hypoleuca*	1A	0.869
F. albicollis vs. *F. hypoleuca*	11	0.816
F. albicollis vs. *F. hypoleuca*	20	0.643
F. albicollis vs. *F. hypoleuca*	Z	0.837

Weak genome-wide correlation between recombination and differentiation early in speciation

Given the longer time since divergence between Collared and Pied Flycatchers, *Ficedula* flycatchers were characterized by a much higher background level of genomic differentiation than capuchinos and wagtails across all chromosomes of interest (mean genome-wide F_{ST}; *S. hypoxantha* vs. *iberaensis*: 0.005, *S. nigrorufa* vs. *S. melanogaster*: 0.008, wagtails: 0.020, *Ficedula* flycatchers: 0.235). We observed a negative correlation between mean F_{ST} and recombination rate in all contrasts (Figure 5.1-5.4). However, this negative relationship was weaker in capuchino seedeaters (Figure 5.1-5.2; *S. hypoxantha*; r = -0.043 to -0.214, *S. iberaensis*: r = -0.102 to -0.205, *S. nigrorufa*: r = -0.172 to -0.306, *S. melanogaster*: r = -0.129 to -0.315) and wagtails (Figure 5.3; *M. a. alba*: r = -0.053 to -0.145, *M. a. personata*: r = -0.089 to -0.142) than in *Ficedula* flycatchers (Figure 5.4; *F. albicollis*: r = -0.379 to -0.561, *F. hypoleuca*: r = -0.347 to -0.54).

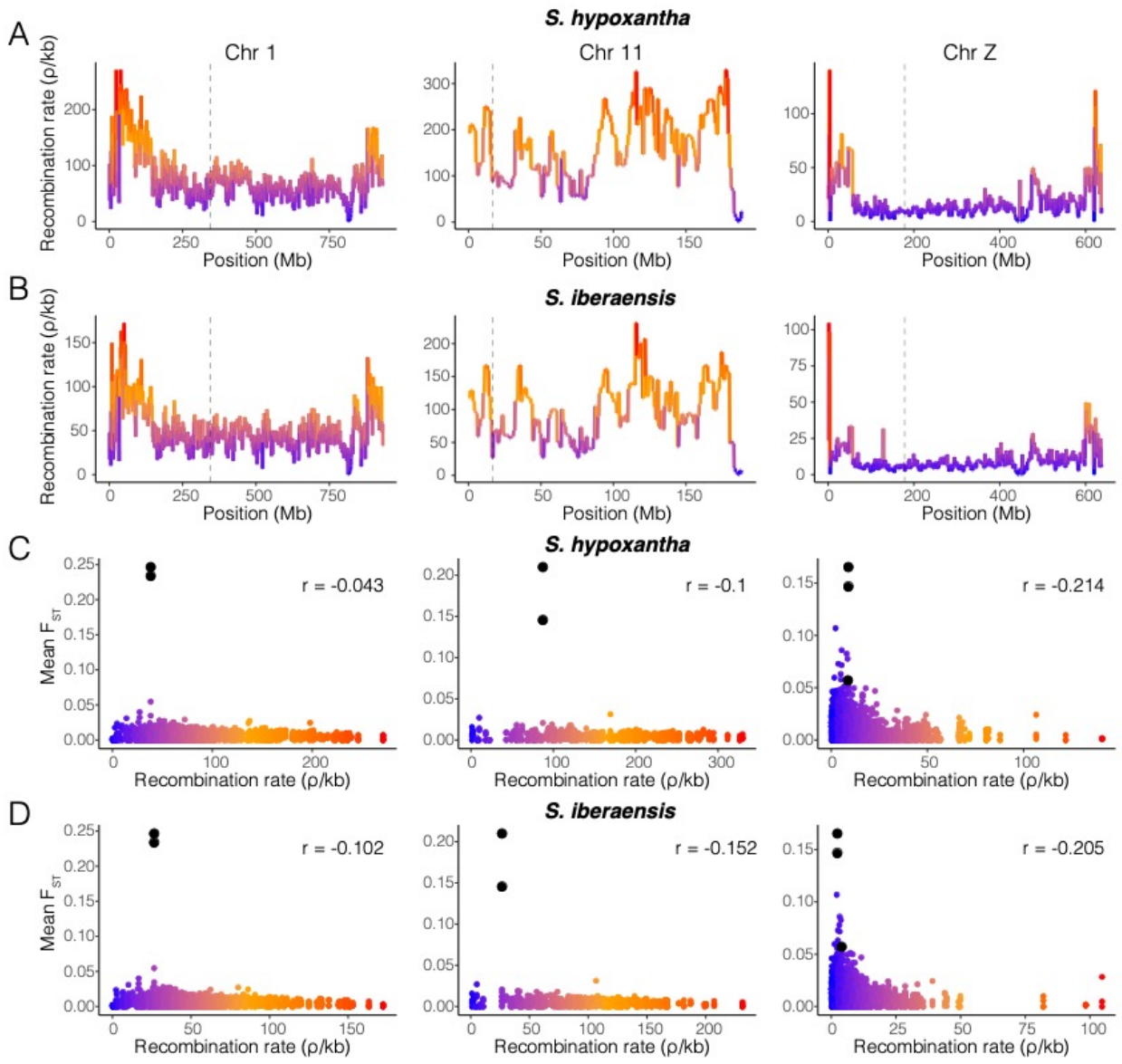


Figure 5.1. (A-B) Fine-scale recombination maps across chromosomes 1, 11, and Z in **(A)** *S. hypoxantha* and **(B)** *S. iberaensis*. Dashed gray lines indicate genomic regions containing divergence peaks. Recombination estimates were averaged over 100-kb windows and are reported in ρ/bp, where ρ is the population-scaled recombination rate ($4N_e r$) and r is the per-generation mutation rate (*i.e.,* the probability of a recombination event occurring during meiosis). **(C-D)** The relationship between mean F_{ST}, calculated over 25-kb windows, and recombination rate in **(C)** *S. hypoxantha* and **(D)** *S. iberaensis*. Windows containing divergence peaks are depicted as black circles and Pearson's correlation coefficient between mean F_{ST} and recombination rate is shown in the upper right corner of each plot. All plots are colored by recombination rate, with regions of low recombination indicated in blue.

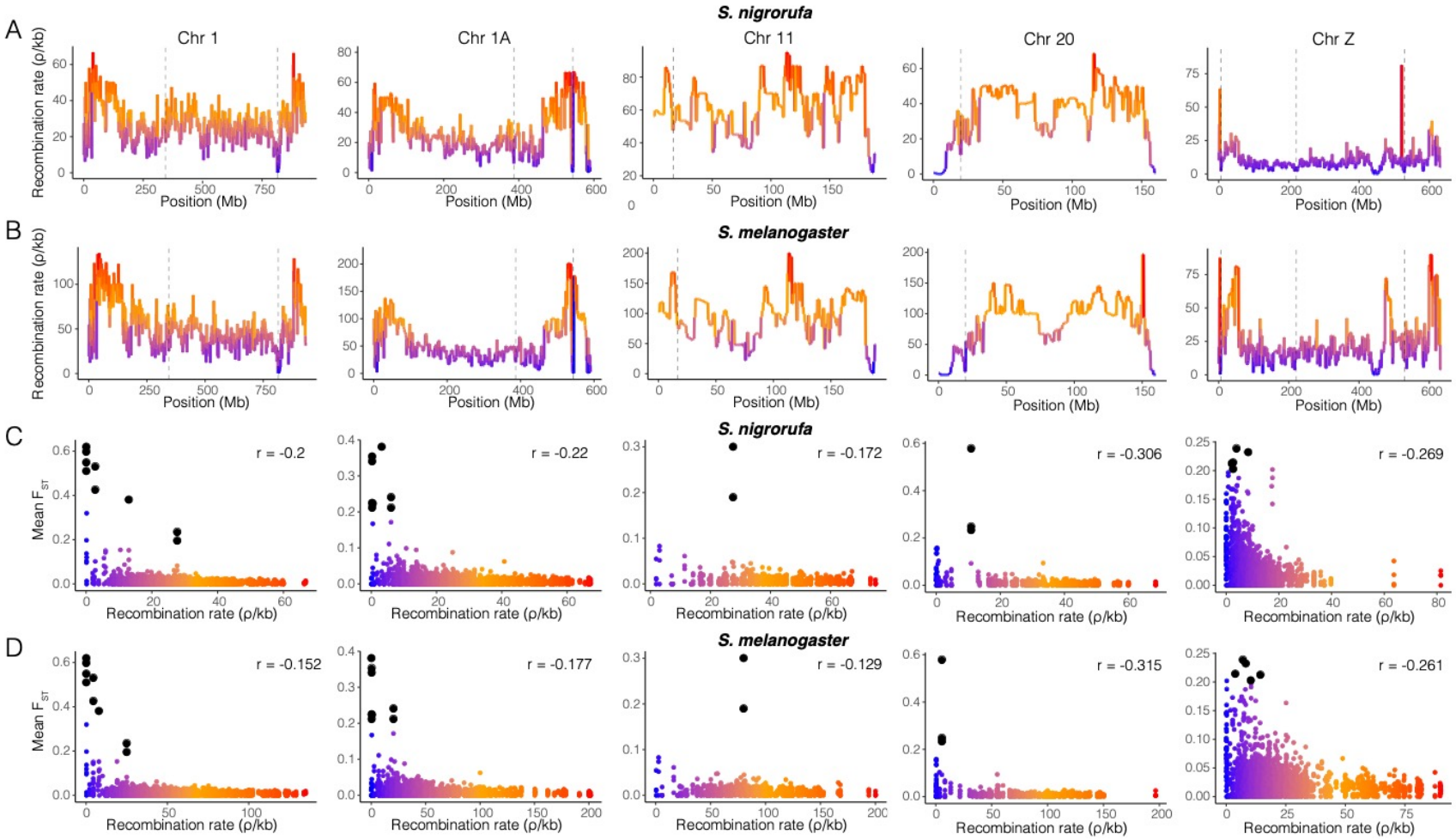


Figure 5.2. (A-B) Fine-scale recombination maps across chromosomes 1, 1A, 11, 20, and Z in **(A)** *S. nigrorufa* and **(B)** *S. melanogaster*. Dashed gray lines indicate genomic regions containing divergence peaks. Details as in Figure 5.1. **(C-D)** The relationship between mean F_{ST}, calculated over 25-kb windows, and recombination rate in **(C)** *S. nigrorufa* and **(D)** *S. melanogaster*. Windows containing divergence peaks are depicted as black circles and Pearson's correlation coefficient between mean F_{ST} and recombination rate is shown in the upper right corner of each plot.

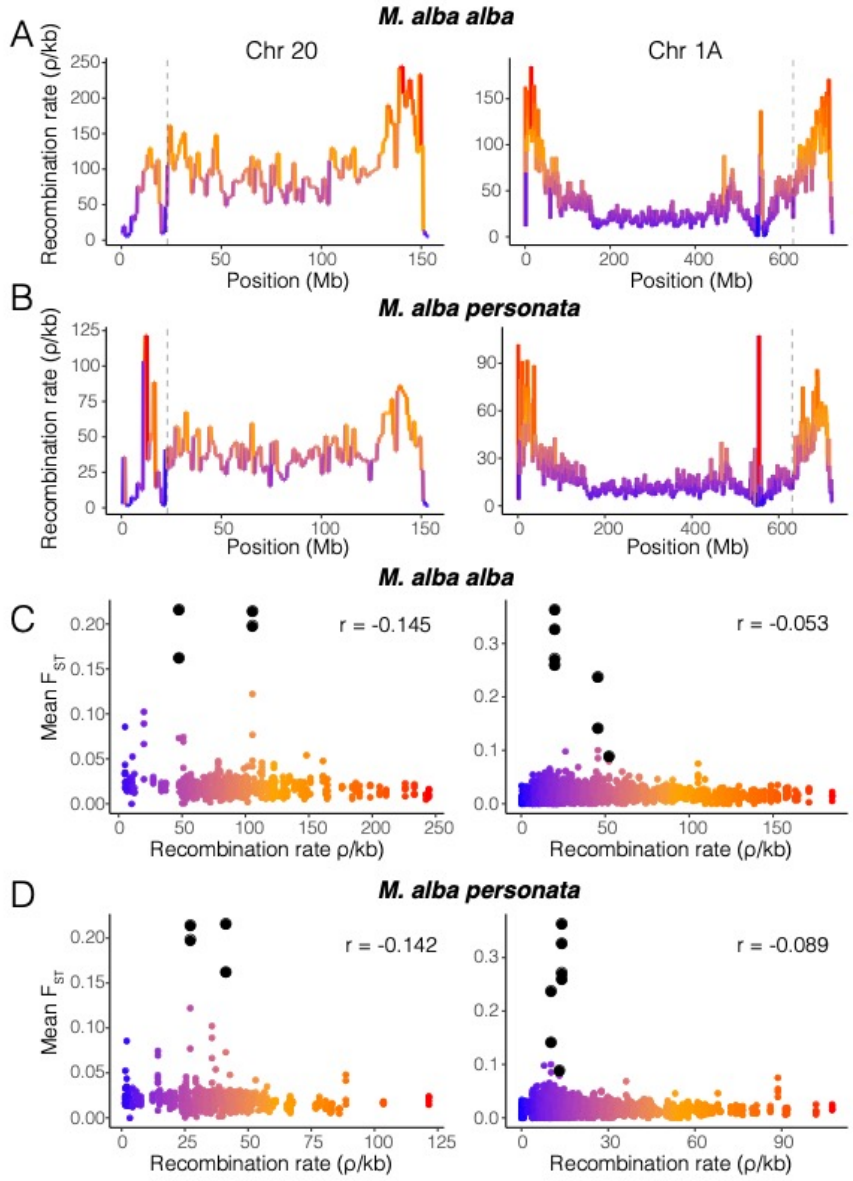


Figure 5.3. (A-B) Fine-scale recombination maps across chromosomes 20 and 1A in **(A)** *M. a. alba* and **(B)** *M. a. personata*. Dashed gray lines indicate genomic regions containing divergence peaks. Details as in Figure 5.1. **(C-D)** The relationship between mean F_{ST}, calculated over 25-kb windows, and recombination rate in **(C)** *M. a. alba* and **(D)** *M. a. personata*. Windows containing divergence peaks are depicted as black circles and Pearson's correlation coefficient between mean F_{ST} and recombination rate is shown in the upper right corner of each plot.

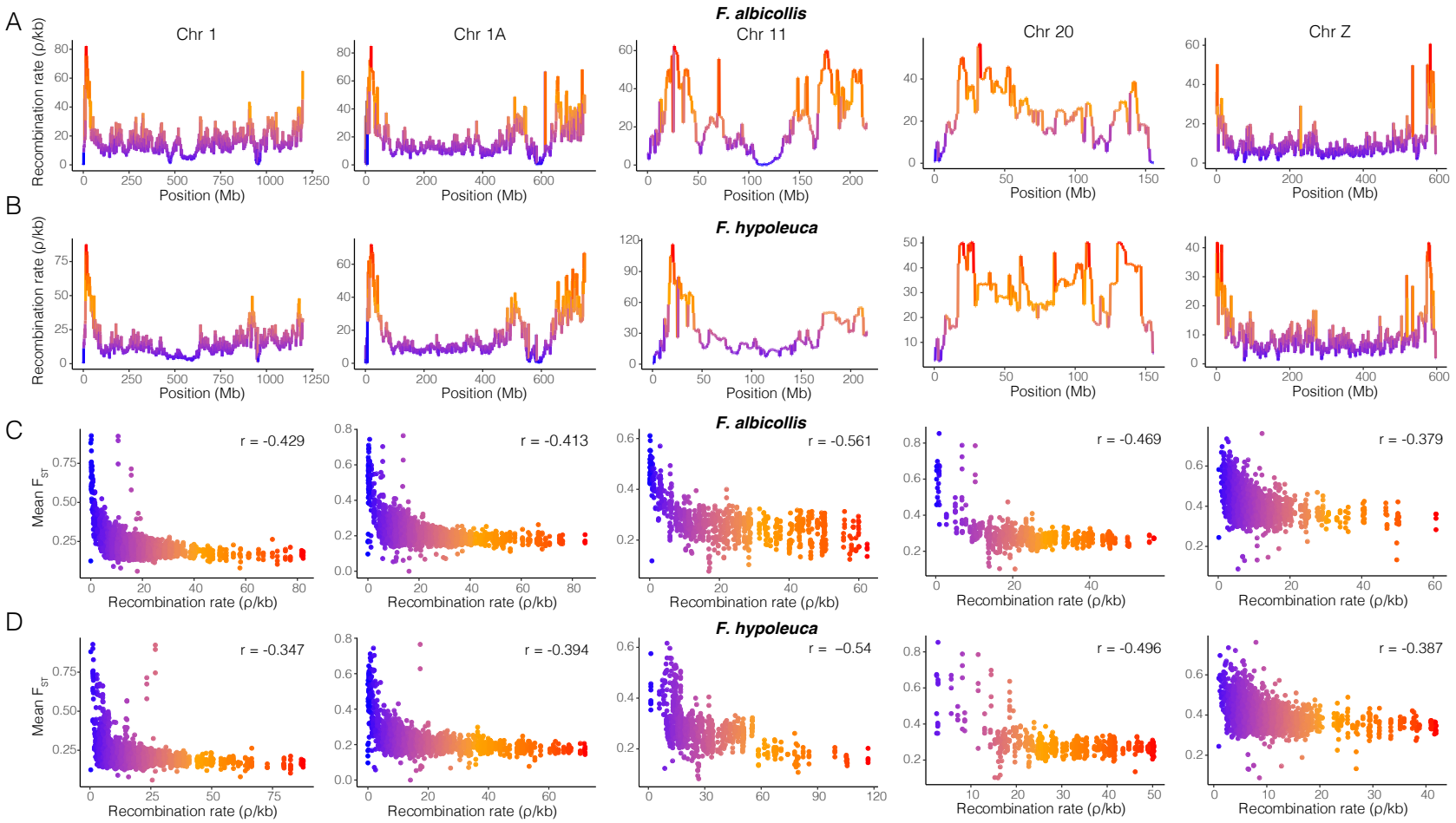


Figure 5.4. (A-B) Fine-scale recombination maps across chromosomes 1, 1A, 11, 20, and Z in **(A)** *F. albicollis* and **(B)** *F. hypoleuca*. Details as in Figure 5.1. **(C-D)** The relationship between mean F_{ST}, calculated over 25-kb windows, and recombination rate in **(C)** *F. albicollis* and **(D)** *F. hypoleuca*. Pearson's correlation coefficient between mean F_{ST} and recombination rate is shown in the upper right corner of each plot.

Many divergence peaks fell in regions of low recombination

Randomization tests revealed that divergence peaks were located in regions of significantly low recombination in 5 of 6 (83%) analyzed chromosomes between *S. hypoxantha* and *S. iberaensis*, 8 of 10 (80%) chromosomes between *S. nigrorufa* and *S. melanogaster*, and 0 of 4 (0%) chromosomes between wagtail subspecies (Table 5.2). However, mean ρ of the divergence peaks was consistently lower than mean ρ of the randomly selected regions in all comparisons (Table 5.2), suggesting that peaks of differentiation tend to fall in regions on the lower end of the recombination spectrum in both capuchinos and wagtails.

Table 5.2. Results of randomization tests to determine if divergence peaks fell in regions with higher or lower recombination rates (ρ) than expected by chance. For each chromosome, 10,000 regions with the same number of windows as the divergence peaks were randomly sampled with replacement. The two-tailed *p* value reports 2 * the proportion of randomly selected regions within the chromosome that had lower recombination rates than that of the divergence peak. Significant results ($p < 0.05$) are shown in bold.

Species	**Chr**	**Mean rho of peaks**	**Mean rho of distribution**	**Two-tailed *p***
S. hypoxantha	1	3834	10491	**0.015**
S. hypoxantha	11	8743	10526	0.612
S. hypoxantha	Z	886	9693	**0**
S. iberaensis	1	2659	6347	**0.040**
S. iberaensis	11	2672	6348	**0.042**
S. iberaensis	Z	291	5969	**0**
S. nigrorufa	1	827	2867	**0**
S. nigrorufa	1A	210	2697	**0**
S. nigrorufa	11	2756	4149	0.121
S. nigrorufa	20	1089	3219	**0.033**
S. nigrorufa	Z	393	1084	**0.001**
S. melanogaster	1	750	5546	**0**
S. melanogaster	1A	538	5652	**0**
S. melanogaster	11	8010	9674	0.487
S. melanogaster	20	551	7901	**0.005**
S. melanogaster	Z	875	2593	**0.002**
M. a. alba	1A	3193	4254	0.457
M. a. alba	20	7661	9651	0.622
M. a. personata	1A	1276	1984	0.234
M. a. personata	20	3420	3853	0.682

5.5 DISCUSSION

Genetic recombination plays a fundamental role in the evolution of sexually reproducing organisms by generating novel allelic combinations that selection can act upon and breaking up associations between loci, thereby allowing selection to remove deleterious mutations that may otherwise accumulate in populations. Nonetheless, we know surprisingly little about the ways in which broad-scale variation in recombination rate contributes to patterns of heterogenous genomic divergence early in the speciation process. With the development of high-throughput genomic sequencing technology, researchers are now able to take advantage of high-density genomic data to estimate fine-scale recombination rates from patterns of linkage disequilibrium in non-model organisms. Using LD-based methods to assess the location of divergence peaks relative to the recombination landscape at various stages of the speciation process, we detected similarities and differences between recent radiations and more divergent taxa.

Within systems, we found that recombination rate was highly correlated among closely related organisms. The conservation of recombination rate across the genome is perhaps unsurprising because genomically similar taxa likely share features that affect patterns of recombination, such as gene density and GC content (Dumont and Payseur 2008; Smukowski and Noor 2011), and chromosomal synteny is high among songbirds (Ellegren 2010). However, in some mammalian taxa, recombination landscapes and the position of hotspots vary substantially among closely related species and subspecies [*e.g.,* between subspecies of mice (Smagulova et al. 2016) and humans and chimpanzees (Auton et al. 2012)]. This discrepancy has been attributed to the rapid evolution of *PRDM9* (Auton et al. 2012; Smagulova et al. 2016), a gene that determines meiotic recombination hotspots in mice and apes but is absent from other vertebrate species, including birds. The observed similarity between the recombination

landscapes of capuchino seedeaters, wagtails, and *Ficedula* flycatchers is therefore consistent with the conservation of hotspot locations reported in birds and other taxa that lack a *PRDM9* gene (Lam and Keeney 2015; Singhal et al. 2015; Kawakami et al. 2017). In capuchino seedeaters, recombination landscapes were remarkably comparable among species despite the fact that particular divergence peaks are present between some species but not others (Campagna et al. 2017). Drastic decreases in diversity (*e.g.,* recurrent selective sweeps) can compromise the ability of LD-based methods to accurately infer recombination rates (Chan et al. 2012), and many of the divergence peaks in capuchino seedeaters are associated with recent species-specific selective sweeps (Hejase et al. 2020). However, LDhelmet is fairly robust to soft selective sweeps, like those observed in capuchinos (Chan et al. 2012; Hejase et al. 2020). Given the similarity between the recombination maps produced for different capuchino species, it is unlikely that positive selection significantly influenced the resulting recombination estimates.

In contrast to Collared and Pied Flycatchers, which exhibited a relatively strong negative correlation between mean F_{ST} and recombination rate across all analyzed chromosomes, we found a much weaker relationship between the two variables in capuchinos and wagtails. Negative associations between F_{ST} and recombination rate have been reported in a variety of taxa (Stephan et al. 1998; Keinan and Reich 2010; Renaut et al. 2013), as the reduction in diversity due to linked selection can extend over larger distances in regions of low recombination, thereby elevating estimates of genomic differentiation. Given that the pattern of heterogeneous genomic divergence in *Ficedula* flycatchers is thought to have largely evolved as a result of background selection and selective sweeps in areas of reduced recombination (Burri et al. 2015), we expected to find a strong negative correlation between mean F_{ST} and recombination rate in Collared and Pied Flycatchers. In capuchino seedeaters and White Wagtails, on the other hand, the few

regions of elevated differentiation between closely related taxa are associated with plumage patterning and are thought to be the result of selective sweeps. We likely did not find a strong relationship between recombination rate and genomic differentiation in the recent radiations at the chromosome level given the short time since divergence between the species and therefore limited influence of background selection in generating the heterogeneous landscape of divergence. In addition, not all divergence peaks in capuchinos and wagtails were located in regions of significantly reduced recombination.

Capuchino seedeaters and White Wagtails are notable for their extensive phenotypic diversity, as both groups have rapidly radiated to form species or subspecies characterized by remarkable variation in plumage patterning. In capuchinos, the reshuffling of standing genetic variation among closely related species is thought to have facilitated phenotypic diversification and rapid speciation (Turbek et al. 2021). High recombination rates in regions containing pigmentations genes could potentially allow the rapid generation of novel plumage patterns, for example through the reshuffling of enhancers that regulate gene expression (Wallbank et al. 2016). However, although we only documented a weak negative correlation between recombination rate and mean F_{ST} in capuchinos and wagtails on a broad scale, we found that a large portion of the identified divergence peaks between capuchino species were located in regions of exceptionally low recombination. In addition, the mean recombination estimates of all divergence peaks in both capuchinos and wagtails fell on the lower end of the distribution of recombination estimates for their respective chromosomes. These results support the findings of previous theoretical and empirical studies suggesting that adaptation and divergence may be favored in regions of low recombination when populations are subject to genetic exchange

(Nachman and Payseur 2012; Marques et al. 2016; Aeschbacher et al. 2017; Han et al. 2017; Samuk et al. 2017; Martin et al. 2019).

During population divergence and speciation, numerous factors, including selection against gene flow, background selection, selective sweeps, and variation in recombination and mutation rates, contribute to heterogeneous genomic differentiation (Ravinet et al. 2017; Semenov et al. 2019). By considering variation in local recombination rate near peaks of divergence produced through known forces in the early stages of the speciation process, we demonstrate that selective sweeps tended to occur in regions of low recombination. Recombination can impede speciation with gene flow by breaking apart favorable allelic combinations and homogenizing divergent populations. Thus, regions responsible for phenotypic differences are less likely to be broken up by gene flow early in divergence if they are located in areas of low recombination, allowing speciation to proceed despite genetic exchange.

CHAPTER 6

CONCLUSION

Recent radiations, such as Darwin's finches, *Heliconius* butterflies, and Lake Victoria cichlids, have long captured the attention of evolutionary biologists and shaped our understanding of diversification and speciation. Given their history of ongoing or recent gene flow, these systems exhibit a relatively low level of background genomic differentiation, allowing researchers to identify genomic regions containing loci that are likely under selection. In addition, they offer the opportunity to study the importance of hybridization in rapid diversification and the ways in which reproductive barriers accumulate over time during the speciation process. By combining fine-scale behavioral analyses with high-throughput genomic sequencing in two recent avian radiations, barn swallows and southern capuchino seedeaters, my dissertation research identified various behavioral and phenotypic traits that contribute to patterns of mate choice and genetic exchange extremely early in the speciation process. In the following section, I summarize the main conclusions of my research and identify the contributions of this work to our understanding of the origin of biodiversity.

Chapter 2 sought to clarify the role of seasonal migration in population divergence and reproductive isolation by organizing hypothesis testing about the ways in which behavioral and phenotypic traits, specifically migratory strategy, may mediate patterns of gene exchange during incipient speciation. In this chapter, I examined five case studies drawn from a variety of taxonomic systems to highlight remaining questions regarding the link between migratory phenotype and reproductive isolation and proposed a conceptual framework to explicitly analyze whether divergent migratory phenotypes lead to genetic differentiation in allopatric scenarios, as

well as pre-mating isolation in migratory divides. This timely paper emphasizes how parallel developments in animal tracking and genomic sequencing technology may be leveraged in concert at migratory divides to track patterns of gene flow as a function of migratory behavior and infer post-mating selection against hybrids. In addition, it highlights a series of outstanding questions regarding our understanding of the contribution of seasonal migration to the speciation process that will hopefully inspire and guide future research in this field.

Chapter 3 applied some of the ideas presented in Chapter 2 to study the association between distinct migratory strategies and genomic differentiation across a hybrid zone between two barn swallow subspecies in Gansu Province, China. By leveraging animal-borne tracking devices, stable isotopes, and high-throughput genomic sequencing, I documented a striking migratory divide that closely tracks population structure across the hybrid zone. Individuals breeding on the western side, with predominately *H. r. rustica ancestry*, had comparatively enriched feather isotope values and overwintered in eastern Africa, while birds breeding on the eastern side, with *H. r. gutturalis ancestry*, had relatively depleted isotope values and migrated to southern India. The two subspecies pursued divergent migratory routes around the high-altitude Karakoram Range and arrived on their breeding grounds over three weeks apart. Collectively, these findings suggest that assortative mating by timing of arrival on the breeding grounds and/or selection against hybrids that inherit intermediate migratory traits may limit interbreeding between the subspecies. In addition, my results support the idea that inhospitable geographic features may have contributed to the diversification of Asian avifauna by influencing migratory patterns. This is the first study to use direct animal-borne tracking methods to examine how migratory songbirds navigate the Qinghai-Tibetan Plateau and similarly inhospitable features in central Asia. More broadly, this study provides insight into the evolutionary consequences of

divergent migratory behavior and the ways in which hostile geographic features may have contributed to the biogeographic patterns of migratory taxa.

Chapter 4 focuses on southern capuchino seedeaters, which comprise one of the most recent and rapid avian radiations on the planet, and combines whole-genome sequencing, citizen science data, and detailed behavioral observations and experiments to evaluate the role of pre-mating isolation in the maintenance of species boundaries between incipient taxa and uncover the genetic basis of the traits that drive it. By analyzing patterns of mate choice between *S. iberaensis* and *S. hypoxantha*, the most abundant capuchino species at the study site, I documented complete assortative mating with regard to both social and extra-pair mates and found evidence for species discrimination based on both plumage patterning and song. In addition, I demonstrated that few genomic variants, encompassing 12 candidate genes, three of which are known to be involved in the regulation of coloration, underlie the novel phenotype of the recently discovered *S. iberaensis* species. Finally, I observed that these genomic regions contain variants that are shared by other members of the capuchino radiation, suggesting that the reshuffling of standing genetic variation led to novel sexual signals that are sufficient to maintain species boundaries extremely early in divergence. This study questions our understanding of the conditions under which incipient speciation may proceed and indicates that variation in access to old genetic variants may be a major factor explaining why some taxa have quickly radiated to form many species while others remain species-poor despite ecological opportunity.

Finally, Chapter 5 examines the role that variation in recombination rate has played in generating peaks of genomic differentiation and the remarkable phenotypic diversity observed in capuchino seedeaters. By generating fine-scale recombination maps for capuchinos and white wagtails, two recent avian radiations in which divergent selection on pigmentation genes likely

generated peaks of differentiation, as well as *Ficedula* flycatchers, in which linked selection has driven divergence patterns, I analyzed the conservation of recombination rates among closely related species and compared the position of divergence peaks relative to regions of exceptionally high and low recombination. I found that recombination landscapes are fairly conserved in all three systems, recombination rate and genomic differentiation are only weakly correlated extremely early in the speciation process, and divergence peaks in capuchinos tend to fall in regions of relatively low recombination. These results indicate that high recombination rates in regions containing pigmentations genes did not facilitate the rapid generation of novel plumage patterns in capuchino seedeaters. On the contrary, genomic regions underlying phenotypic differences in capuchinos were likely able to persist in the face of gene flow due to their location in areas of low recombination, facilitating speciation in sympatry.

Collectively, these studies highlight the power of combining field-based behavioral experiments with genomic analyses to gain a more thorough picture of the processes maintaining species boundaries among genomically similar taxa. By identifying the genomic basis of phenotypic differences, evaluating the importance of those phenotypic traits for patterns of genetic exchange, and examining the origin of genomic variants that differentiate closely related species, we can begin to fully understand the processes that generate and maintain the incredible biodiversity that we observe in the natural world.

REFERENCES

Abolins-Abols, M., E. Kornobis, P. Ribeca, K. Wakamatsu, M. P. Peterson, E. D. Ketterson, and B. Milá. 2018. Differential gene regulation underlies variation in melanic plumage coloration in the dark-eyed junco (*Junco hyemalis*). Mol. Ecol. 27:4501–4515.

Aeschbacher, S., J. P. Selby, J. H. Willis, and G. Coop. 2017. Population-genomic inference of the strength and timing of selection against gene flow. Proc. Natl. Acad. Sci. 114:7061–7066.

Albert, V., B. Jónsson, and L. Bernatchez. 2006. Natural hybrids in Atlantic eels (*Anguilla anguilla*, *A. rostrata*): evidence for successful reproduction and fluctuating abundance in space and time. Mol. Ecol. 15:1903–1916.

Alexander, D. H., J. Novembre, and K. Lange. 2009. Fast model-based estimation of ancestry in unrelated individuals. Genome Res. 19:1655–1664.

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

Alvarado, A. H., T. L. Fuller, and T. B. Smith. 2014. Integrative tracking methods elucidate the evolutionary dynamics of a migratory divide. Ecol. Evol. 4:3456–3469.

Andersson, M. B. 1994. Sexual selection. Princeton University Press, Princeton.

Andrews, K. R., J. M. Good, M. R. Miller, G. Luikart, and P. A. Hohenlohe. 2016. Harnessing the power of RADseq for ecological and evolutionary genomics. Nat. Rev. Genet. 17:81–92.

Andrews, S. 2010. FastQC: a quality control tool for high throughput sequence data.

Aristotle. *c*. 350 BC. A History of Animals, Book IX, Part 49B (Thompson, D.W., transl.). http://classics.mit.edu/Aristotle/ history_anim.mb.txt

Auton, A., A. Fledel-Alon, S. Pfeifer, O. Venn, L. Ségurel, T. Street, E. M. Leffler, R. Bowden, I. Aneas, J. Broxholme, P. Humburg, Z. Iqbal, G. Lunter, J. Maller, R. D. Hernandez, C. Melton, A. Venkat, M. A. Nobrega, R. Bontrop, S. Myers, P. Donnelly, M. Przeworski, and G. McVean. 2012. A Fine-Scale Chimpanzee Genetic Map from Population Sequencing. Science. 336:193–198.

Baker, C. S., S. R. Palumbi, R. H. Lambertsen, M. T. Weinrich, J. Calambokidis, and S. J. O'Brien. 1990. Influence of seasonal migration on geographic distribution of mitochondrial DNA haplotypes in humpback whales. Nature 344:238–240.

Baker, M. C. 1991. Response of Male Indigo and Lazuli Buntings and Their Hybrids To Song Playback in Allopatric and Sympatric Populations. Behaviour 119:225–242.

Baker, M. C., and A. E. M. Baker. 1990. Reproductive behavior of female buntings: isolating mechanisms in a hybridizing pair of species. Evolution. 44:332–338.

Balakrishnan, C. N., and M. D. Sorenson. 2006. Song discrimination suggests premating isolation among sympatric indigobird species and host races. Behav. Ecol. 17:473–478.

Bates, D., M. Mächler, B. Bolker, and S. Walker. 2015. Fitting Linear Mixed-Effects Models Using lme4. J. Stat. Software; Vol 1.

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.

Benites, P., L. Campagna, and P. L. Tubaro. 2015. Song-based species discrimination in a rapid Neotropical radiation of grassland seedeaters. J. Avian Biol. 46:55–62.

Benites, P., M. D. Eaton, D. A. Lijtmaer, S. C. Lougheed, and P. L. Tubaro. 2010. Analysis from avian visual perspective reveals plumage colour differences among females of capuchino

seedeaters (*Sporophila*). J. Avian Biol. 41:597–602.

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. Mol. Ecol. 18:3087–3096.

Berner, D., and M. Roesti. 2017. Genomics of adaptive divergence with chromosome-scale heterogeneity in crossover rate. Mol. Ecol. 26:6351–6369.

Berthold, P. 1991. Orientation in birds. Birkhäuser Verlag, Basel.

Berthold, P., A. J. Helbig, G. Mohr, and U. Querner. 1992. Rapid microevolution of migratory behaviour in a wild bird species. Nature 360:668–670.

BirdLife International. 2020. IUCN Red List for birds.

Bivand, R., and C. Rundel. 2017. rgeos: interface to geometry engine-open source (GEOS). R Package version 0.3-23.

Bobek, M., R. Hampl, L. Peške, F. Pojer, J. Šimek, and S. Bureš. 2008. African Odyssey project – satellite tracking of black storks *Ciconia nigra* breeding at a migratory divide. J. Avian Biol. 39:500–506.

Bolger, A. M., M. Lohse, and B. Usadel. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120.

Booker, T. R., S. Yeaman, and M. C. Whitlock. 2020. Variation in recombination rate affects detection of outliers in genome scans under neutrality. Mol. Ecol. 29:4274–4279.

Bowen, B., J. C. Avise, J. I. Richardson, A. B. Meylan, D. Margaritoulis, and S. R. Hopkins-Murphy. 1993. Population Structure of Loggerhead Turtles (*Caretta caretta*) in the

Northwestern Atlantic Ocean and Mediterranean Sea. Conserv. Biol. 7:834–844.

Bowen, G. J., and J. Revenaugh. 2003. Interpolating the isotopic composition of modern meteoric precipitation. Water Resour. Res. 39.

Broad Institute, Picard Toolkit. 2018. http://broadinstitute.github.io/picard

Bridge, E. S., K. Thorup, M. S. Bowlin, P. B. Chilson, R. H. Diehl, R. W. Fléron, P. Hartl, R. Kays, J. F. Kelly, W. D. Robinson, and M. Wikelski. 2011. Technology on the Move: Recent and Forthcoming Innovations for Tracking Migratory Birds. Bioscience 61:689–698.

Brooks, M. E., K. Kristensen, K. J. van Benthem, A. Magnusson, C. W. Berg, A. Nielsen, H. J. Skaug, M. Machler, and B. M. Bolker. 2017. glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. R J. 9:378–400.

Browning, S. R., and B. L. Browning. 2007. Rapid and Accurate Haplotype Phasing and Missing-Data Inference for Whole-Genome Association Studies By Use of Localized Haplotype Clustering. Am. J. Hum. Genet. 81:1084–1097.

Burns, K. J., A. J. Shultz, P. O. Title, N. A. Mason, F. K. Barker, J. Klicka, S. M. Lanyon, and I. J. Lovette. 2014. Phylogenetics and diversification of tanagers (Passeriformes: Thraupidae), the largest radiation of Neotropical songbirds. Mol. Phylogenet. Evol. 75:41–77.

Burri, R. 2017. Interpreting differentiation landscapes in the light of long-term linked selection. Evol. Lett. 1:118–131.

Burri, R., A. Nater, T. Kawakami, C. F. Mugal, P. I. Olason, L. Smeds, A. Suh, L. Dutoit, S. Bureš, L. Z. Garamszegi, S. Hogner, J. Moreno, A. Qvarnström, M. Ružić, S.-A. Sæther, G.-P. Sætre, J. Török, and H. Ellegren. 2015. Linked selection and recombination rate variation drive the evolution of the genomic landscape of differentiation across the

speciation continuum of *Ficedula* flycatchers. Genome Res. 25:1656–1665.

Burri, R. 2017. Interpreting differentiation landscapes in the light of long-term linked selection. Evol. Lett. 1:118–131.

Butlin, R., A. Debelle, C. Kerth, R. Snook, L. Beukeboom, R. Castillo, W. Diao, M. Maan, S. Paolucci, F. Weissing, and A. Hoikkala. 2012. What do we need to know about speciation? Trends Ecol. Evol. 27:27–39.

Butlin, R. K. 2005. Recombination and speciation. Mol. Ecol. 14:2621–2635.

Butlin, R. K., S. Maria, C. Grégory, J. Benjamin, A. Carl, C. Armando, C. J. A., G. Juan, G. J. W., H. Johan, K. Petri, M. Mónica, P. Marina, Q. Humberto, J. Kerstin, and R. Emilio. 2013. Parallel Evolution of Local Adaptation and Reproductive Isolation in the Face of Gene Flow. Evolution. 68:935–949.

Campagna, L., P. Benites, S. C. Lougheed, D. A. Lijtmaer, A. S. Di Giacomo, M. D. Eaton, and P. L. Tubaro. 2012. Rapid phenotypic evolution during incipient speciation in a continental avian radiation. Proc. R. Soc. B Biol. Sci. 279:1847–1856.

Campagna, L., I. Gronau, L. F. Silveira, A. Siepel, and I. J. Lovette. 2015. Distinguishing noise from signal in patterns of genomic divergence in a highly polymorphic avian radiation. Mol. Ecol. 24:4238–4251.

Campagna, L., D. A. Lijtmaer, K. C. R. Kerr, A. S. Barreira, P. D. N. Hebert, S. C. Lougheed, and P. L. Tubaro. 2010. DNA barcodes provide new evidence of a recent radiation in the genus *Sporophila* (Aves: Passeriformes). Mol. Ecol. Resour. 10:449–458.

Campagna, L., K. G. McCracken, and I. J. Lovette. 2019. Gradual evolution towards flightlessness in steamer ducks. Evolution. 73:1916–1926.

Campagna, L., M. Repenning, L. F. Silveira, C. S. Fontana, P. L. Tubaro, and I. J. Lovette. 2017.

Repeated divergent selection on pigmentation genes in a rapid finch radiation. Sci. Adv. 3:e1602404.

Campagna, L., P. Rodriguez, and J. C. Mazzulla. 2018. Transgressive phenotypes and evidence of weak postzygotic isolation in F1 hybrids between closely related capuchino seedeaters. PLoS One 13:e0199113.

Campagna, L., L. F. Silveira, P. L. Tubaro, and S. C. Lougheed. 2013. Identifying the Sister Species to the Rapid Capuchino Seedeater Radiation (Passeriformes: *Sporophila*). Auk 130:645–655.

Catchen, J., P. A. Hohenlohe, S. Bassham, A. Amores, and W. A. Cresko. 2013. Stacks: an analysis tool set for population genomics. Mol. Ecol. 22:3124–3140.

Chamberlain, C. P., S. Bensch, X. Feng, S. Åkesson, 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. Proc. R. Soc. London. Ser. B Biol. Sci. 267:43–48.

Chan, A. H., P. A. Jenkins, and Y. S. Song. 2012. Genome-Wide Fine-Scale Recombination Rate Variation in *Drosophila melanogaster*. PLOS Genet. 8:e1003090.

Charlesworth, B., M. T. Morgan, and D. Charlesworth. 1993. The effect of deleterious mutations on neutral molecular variation. Genetics 134:1289–1303.

Chifman, J., and L. Kubatko. 2014. Quartet Inference from SNP Data Under the Coalescent Model. Bioinformatics 30:3317–3324.

Collar, N. J. 1992. Threatened birds of the Americas. Smithsonian Institution Press in cooperation with International Council for Bird Preservation.

Collins, S. A., S. R. de Kort, J. Pérez-Tris, and J. Luis Tellería. 2009. Migration strategy and

divergent sexual selection on bird song. Proc. R. Soc. B Biol. Sci. 276:585–590.

Comeron, J. M., A. Williford, and R. M. Kliman. 2008. The Hill–Robertson effect: evolutionary consequences of weak selection and linkage in finite populations. Heredity 100:19–31.

Cooper, E. A., and J. A. C. Uy. 2017. Genomic evidence for convergent evolution of a key trait underlying divergence in island birds. Mol. Ecol. 26:3760–3774.

Coyne, J. A., and H. A. Orr. 1989. Patterns of speciation in *Drosophila*. Evolution. 43:362–381.

Coyne, J., and H. Orr. 2004. Speciation. Sinauer Associates, Sunderland.

Craig, J. K., and C. J. Foote. 2001. Countergradient variation and secondary sexual color: Phenotypic convergence promotes genetic divergence in carotenoid use between sympatric anadromous and nonanadromous morphs of Sockeye Salmon (*Oncorhynchus nerka*). Evolution. 55:380–391.

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

Cuthill, I. C., J. C. Partridge, A. T. D. Bennett, S. C. Church, N. S. Hart, and S. Hunt. 2000. Ultraviolet Vision in Birds. Adv. Study Behav. 29:159–214.

da Silva, J. M. C. 1999. Seasonal movements and conservation of seedeaters of the genus *Sporophila* in South America. Stud. Avian Biol. 19:272–280.

Danecek, P., A. Auton, G. Abecasis, C. A. Albers, E. Banks, M. A. DePristo, R. E. Handsaker, G. Lunter, G. T. Marth, S. T. Sherry, G. McVean, R. Durbin, and 1000 Genomes Project Analysis Group. 2011. The variant call format and VCFtools. Bioinformatics 27:2156–2158.

Davis, S. E., M. Maftei, and M. L. Mallory. 2016. Migratory Connectivity at High Latitudes: Sabine's Gulls (*Xema sabini*) from a Colony in the Canadian High Arctic Migrate to

Different Oceans. PLoS One 11:e0166043.

Dawson, T. E., S. Mambelli, A. H. Plamboeck, P. H. Templer, and K. P. Tu. 2002. Stable Isotopes in Plant Ecology. Annu. Rev. Ecol. Syst. 33:507–559.

del Hoyo, J., A. Elliott, J. Sargatal, D. A. Christie, and G. Kirwan. 2020. Handbook of the Birds of the World Alive.

Delaneau, O., J.-F. Zagury, and J. Marchini. 2013. Improved whole-chromosome phasing for disease and population genetic studies. Nat. Methods 10:5–6.

Delany, S., C. Williams, C. Sulston, J. Norton, and D. Garbutt. 2017. Passerine Migration across the Himalayas. Pp. 58–81 *in* H. H. T. Prins and T. Namgail, eds. Bird migration across the Himalayas: wetland functioning amidst mountains and glaciers. Cambridge University Press, Cambridge.

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. Proc. R. Soc. B Biol. Sci. 279:4582–4589.

Delmore, K. E., S. Hübner, N. C. Kane, R. Schuster, R. L. Andrew, F. Câmara, R. Guigó, and D. E. Irwin. 2015a. Genomic analysis of a migratory divide reveals candidate genes for migration and implicates selective sweeps in generating islands of differentiation. Mol. Ecol. 24:1873–1888.

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

Delmore, K. E., H. L. Kenyon, R. R. Germain, and D. E. Irwin. 2015b. Phenotypic divergence during speciation is inversely associated with differences in seasonal migration. Proc. R. Soc. B Biol. Sci. 282:20151921.

Delmore, K. E., and M. Liedvogel. 2016. Investigating factors that generate and maintain variation in migratory orientation: A primer for recent and future work. Front. Behav. Neurosci. 10:3.

Delmore, K. E., D. P. L. Toews, R. R. Germain, G. L. Owens, and D. E. Irwin. 2016. The Genetics of Seasonal Migration and Plumage Color. Curr. Biol. 26:2167–2173.

Di Giacomo, A. G., and S. F. Krapovickas. 2005. Historia natural y paisaje de la Reserva El Bagual, provincia de Formosa, Argentina. Temas Nat. y Conserv. 4:1–592.

Di Giacomo, A. S., and M. S. Abril. 2005. Áreas importantes para la conservación de las aves en la Argentina: sitios prioritarios para la conservación de la biodiversidad. Aves Argentinas/Asociación Ornitológica del Plata.

Di Giacomo, A. S., and C. Kopuchian. 2016. Una nueva especie de capuchino (*Sporophila*: Thraupidae) de los Esteros del Iberá, Corrientes, Argentina. Nuestras Aves 61:3–5.

Dingle, H. 2014. Migration: the biology of life on the move. Oxford University Press, New York.

Dodson, J. J., N. Aubin-Horth, V. Thériault, and D. J. Páez. 2013. The evolutionary ecology of alternative migratory tactics in salmonid fishes. Biol. Rev. 88:602–625.

Domyan, E. T., M. W. Guernsey, Z. Kronenberg, S. Krishnan, R. E. Boissy, A. I. Vickrey, C. Rodgers, P. Cassidy, S. A. Leachman, J. W. Fondon, M. Yandell, and M. D. Shapiro. 2014. Epistatic and Combinatorial Effects of Pigmentary Gene Mutations in the Domestic Pigeon. Curr. Biol. 24:459–464.

Dong, X., T. Pan, X. Kang, Y. Zhang, X. Sun, and L. Qian. 2016. Complete mitochondrial genome of *Motacilla alba* and implications for Motacillidae taxonomy. Mitochondrial DNA Part A 27:4675–4676.

Dumont, B. L., and B. A. Payseur. 2008. Evolution of the genomic rate of recombination in mammals. Evolution. 62:276–294.

eBird Basic Dataset. Version: EBD_relJan-2020. Cornell Lab of Ornithology, Ithaca, New York. Jan 2020.

Ellegren, H. 2009. Genomic evidence for a large-Z effect. Proc. R. Soc. B Biol. Sci. 276:361–366.

Ellegren, H. 2010. Evolutionary stasis: the stable chromosomes of birds. Trends Ecol. Evol. 25:283–291.

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.

Farrington, J. D. 2016. A survey of the autumn 2009 and spring 2010 bird migrations at Lhasa, Tibet Autonomous Region, China. Forktail 32:14–25.

Feder, J. L., S. P. Egan, and P. Nosil. 2018. The genomics of speciation-with-gene-flow. Trends Genet. 28:342–350.

Foote, C. J., and P. A. Larkin. 1988. The role of male choice in the assortative mating of anadromous and non-anadromous sockeye salmon (*Oncorhynchus nerka*). Behaviour 106:43–62.

Forbes, A. A., S. N. Devine, A. C. Hippee, E. S. Tvedte, A. K. G. Ward, H. A. Widmayer, and C. J. Wilson. 2017. Revisiting the particular role of host shifts in initiating insect speciation. Evolution. 71:1126–1137.

Formenti, G., M. Chiara, L. Poveda, K.-J. Francoijs, A. Bonisoli-Alquati, L. Canova, L. Gianfranceschi, D. S. Horner, and N. Saino. 2018. SMRT long reads and Direct Label and

Stain optical maps allow the generation of a high-quality genome assembly for the European barn swallow (*Hirundo rustica rustica*). Gigascience 8.

Fraga, R. M. 2001. The avifauna of Estancia San Juan Poriahú, Iberá Marshes, Argentina: checklist and some natural history notes. Cotinga 16:81–86.

Franchini, P., I. Irisarri, A. Fudickar, A. Schmidt, A. Meyer, M. Wikelski, and J. Partecke. 2017. Animal tracking meets migration genomics: transcriptomic analysis of a partially migratory bird species. Mol. Ecol. 26:3204–3216.

Freed, D., R. Aldana, J. A. Weber, and J. S. Edwards. 2017. The Sentieon Genomics Tools-A fast and accurate solution to variant calling from next-generation sequence data. BioRxiv 115717.

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. Proc. Natl. Acad. Sci. 104:18589–18594.

Fry, B., W. Brand, F. J. Mersch, K. Tholke, and R. Garritt. 1992. Automated analysis system for coupled .delta.13C and .delta.15N measurements. Anal. Chem. 64:288–291.

García-Alcalde, F., K. Okonechnikov, J. Carbonell, L. M. Cruz, S. Götz, S. Tarazona, J. Dopazo, T. F. Meyer, and A. Conesa. 2012. Qualimap: evaluating next-generation sequencing alignment data. Bioinformatics 28:2678–2679.

Giraudo, A. R., M. A. Ordano, M. Chatellenaz, E. R. Krauczuk, A. H. Beltzer, C. Saibene, A. S. Di Giacomo, and J. Alonso. 2003. Aves de los esteros del Iberá (Corrientes, Argentina): Patrones de diversidad, historia natural y perspectivas de conservación. Fauna del Iberá. Editorial. Univ. la Univ. Nac. del Noreste, Corrientes.

Gompert, Z., E. G. Mandeville, and C. A. Buerkle. 2017. Analysis of population genomic data

from hybrid zones. Annu. Rev. Ecol. Evol. Syst. 48.

Grabherr, M. G., P. Russell, M. Meyer, E. Mauceli, J. Alföldi, F. Di Palma, and K. Lindblad-Toh. 2010. Genome-wide synteny through highly sensitive sequence alignment: Satsuma. Bioinformatics 26:1145–1151.

Grant, B. R., and P. R. Grant. 2002a. Lack of Premating Isolation at the Base of a Phylogenetic Tree. Am. Nat. 160:1–19.

Grant, B. R., and P. R. Grant. 2002b. Simulating secondary contact in allopatric speciation: an empirical test of premating isolation. Biol. J. Linn. Soc. 76:545–556.

Grant, P. R., and B. R. Grant. 2018. Role of sexual imprinting in assortative mating and premating isolation in Darwin's finches. Proc. Natl. Acad. Sci. 115:E10879–E10887.

Griffith, S. C., I. P. F. Owens, and K. A. Thuman. 2002. Extra pair paternity in birds: a review of interspecific variation and adaptive function. Mol. Ecol. 11:2195–2212.

Gronau, I., M. J. Hubisz, B. Gulko, C. G. Danko, and A. Siepel. 2011. Bayesian inference of ancient human demography from individual genome sequences. Nat. Genet. 43:1031.

Hahn, C., L. Bachmann, and B. Chevreux. 2013. Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads—a baiting and iterative mapping approach. Nucleic Acids Res. 41:e129–e129.

Han, F., S. Lamichhaney, B. R. Grant, P. R. Grant, L. Andersson, and M. T. Webster. 2017. Gene flow, ancient polymorphism, and ecological adaptation shape the genomic landscape of divergence among Darwin's finches. Genome Res. 27:1004–1015.

Harris, R. B., P. Alström, A. Ödeen, and A. D. Leaché. 2018. Discordance between genomic divergence and phenotypic variation in a rapidly evolving avian genus (*Motacilla*). Mol. Phylogenet. Evol. 120:183–195.

Hebert, P. D. N., S. Ratnasingham, and J. R. de Waard. 2003. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. Proc. R. Soc. London. Ser. B Biol. Sci. 270:S96–S99.

Hecht, B. C., N. R. Campbell, D. E. Holecek, and S. R. Narum. 2013. Genome-wide association reveals genetic basis for the propensity to migrate in wild populations of rainbow and steelhead trout. Mol. Ecol. 22:3061–3076.

Hejase, H. A., A. Salman-Minkov, L. Campagna, M. J. Hubisz, I. J. Lovette, I. Gronau, and A. Siepel. 2020. Genomic islands of differentiation in a rapid avian radiation have been driven by recent selective sweeps. Proc. Natl. Acad. Sci. 117:30554–30565.

Helbig, A. 1996. Genetic basis, mode of inheritance and evolutionary changes of migratory directions in palaearctic warblers (Aves: Sylviidae). J. Exp. Biol. 199:49–55.

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

Heng, L., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, and 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25:2078–2079.

Higgie, M., and M. W. Blows. 2007. Are traits that experience reinforcement also under sexual selection? Am. Nat. 170:409–420.

Hill, W. G., and A. Robertson. 1966. The effect of linkage on limits to artificial selection. Genet. Res. 8:269–294.

Hobbie, E. A., and R. A. Werner. 2004. Intramolecular, compound-specific, and bulk carbon isotope patterns in C3 and C4 plants: a review and synthesis. New Phytol. 161:371–385.

Hobson, K. A. 1999. Tracing origins and migration of wildlife using stable isotopes: a review. Oecologia 120:314–326.

Hoskin, C. J., M. Higgie, K. R. McDonald, and C. Moritz. 2005. Reinforcement drives rapid allopatric speciation. Nature 437:1353.

Howard, D. J. 1993. Reinforcement: origin, dynamics, and fate of an evolutionary hypothesis. Pp. 46–69 *in* R. G. Harrison, ed. Hybrid zones and the evolutionary process. Oxford University Press, New York.

International Chicken Genomic Sequencing Consortium. 2004. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. Nature 432:695–716.

Irwin, D. E. 2000. Song variation in an avian ring species. Evolution. 54:998–1010.

Irwin, D. E. 2009. Speciation: New migratory direction provides route toward divergence. Curr. Biol. 19:R1111–R1113.

Irwin, D. E. 2020. Assortative mating in hybrid zones is remarkably ineffective in promoting speciation. Am. Nat. 195:E150-E167.

Irwin, D. E., S. Bensch, and T. D. Price. 2001. Speciation in a ring. Nature 409:333–337.

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. 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. S. 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. Mol. Ecol. 20:3102–3115.

Jacobsen, M. W., J. M. Pujolar, L. Bernatchez, K. Munch, J. Jian, Y. Niu, and M. M. Hansen.

2014. Genomic footprints of speciation in Atlantic eels (*Anguilla anguilla* and *A. rostrata*). Mol. Ecol. 23:4785–4798.

Jain, C., A. Dilthey, S. Koren, S. Aluru, and A. M. Phillippy. 2017. A Fast Approximate Algorithm for Mapping Long Reads to Large Reference Databases. J. Comput. Biol. 25:66–81.

Jain, C., S. Koren, A. Dilthey, A. M. Phillippy, and S. Aluru. 2018. A fast adaptive algorithm for computing whole-genome homology maps. Bioinformatics 34:i748–i756.

Jorgensen, S. J., C. A. Reeb, T. K. Chapple, S. Anderson, C. Perle, S. R. Van Sommeran, C. Fritz-Cope, A. C. Brown, A. P. Klimley, and B. A. Block. 2010. Philopatry and migration of Pacific white sharks. Proc. R. Soc. B Biol. Sci. 277:679–688.

Kalinowski, S. T., M. L. Taper, and T. C. Marshall. 2007. Revising how the computer program cervus accommodates genotyping error increases success in paternity assignment. Mol. Ecol. 16:1099–1106.

Karpestam, E., L. Wennersten, and A. Forsman. 2012. Matching habitat choice by experimentally mismatched phenotypes. Evol. Ecol. 26:893–907.

Kawakami, T., C. F. Mugal, A. Suh, A. Nater, R. Burri, L. Smeds, and H. Ellegren. 2017. Whole-genome patterns of linkage disequilibrium across flycatcher populations clarify the causes and consequences of fine-scale recombination rate variation in birds. Mol. Ecol. 26:4158–4172.

Kawakami, T., L. Smeds, N. Backström, A. Husby, A. Qvarnström, C. F. Mugal, P. Olason, and H. Ellegren. 2014. A high-density linkage map enables a second-generation collared flycatcher genome assembly and reveals the patterns of avian recombination rate variation and chromosomal evolution. Mol. Ecol. 23:4035–4058.

Kays, R., M. C. Crofoot, W. Jetz, and M. Wikelski. 2015. Terrestrial animal tracking as an eye on life and planet. Science. 348:aaa2478.

Keagy, J., L. Lettieri, and J. W. Boughman. 2016. Male competition fitness landscapes predict both forward and reverse speciation. Ecol. Lett. 19:71–80.

Kearse, M., R. Moir, A. Wilson, S. Stones-Havas, M. Cheung, S. Sturrock, S. Buxton, A. Cooper, S. Markowitz, C. Duran, T. Thierer, B. Ashton, P. Meintjes, and A. Drummond. 2012. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28:1647–1649.

Keinan, A., and D. Reich. 2010. Human Population Differentiation Is Strongly Correlated with Local Recombination Rate. PLOS Genet. 6:e1000886.

Kenny, E. E., N. J. Timpson, M. Sikora, M.-C. Yee, A. Moreno-Estrada, C. Eng, S. Huntsman, E. G. Burchard, M. Stoneking, C. D. Bustamante, and S. Myles. 2012. Melanesian Blond Hair Is Caused by an Amino Acid Change in TYRP1. Science. 336:554–554.

Kent, W. J. 2002. BLAT—The BLAST-Like Alignment Tool. Genome Res. 12:656–664.

Kent, W. J., C. W. Sugnet, T. S. Furey, K. M. Roskin, T. H. Pringle, A. M. Zahler, and and D. Haussler. 2002. The Human Genome Browser at UCSC. Genome Res. 12:996–1006.

Kerr, K. C. R., D. A. Lijtmaer, A. S. Barreira, P. D. N. Hebert, and P. L. Tubaro. 2009. Probing Evolutionary Patterns in Neotropical Birds through DNA Barcodes. PLoS One 4:e4379.

Klaassen, H., Y. Wang, K. Adamski, N. Rohner, and J. E. Kowalko. 2018. CRISPR mutagenesis confirms the role of oca2 in melanin pigmentation in *Astyanax mexicanus*. Dev. Biol. 441:313–318.

Kopp, M., M. R. Servedio, T. C. Mendelson, R. J. Safran, R. L. Rodríguez, M. E. Hauber, E. C. Scordato, L. B. Symes, C. N. Balakrishnan, D. M. Zonana, and G. S. van Doorn. 2017.

Mechanisms of Assortative Mating in Speciation with Gene Flow: Connecting Theory and Empirical Research. Am. Nat. 191:1–20.

Kozlov, A. M., D. Darriba, T. Flouri, B. Morel, and A. Stamatakis. 2019. RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. Bioinformatics 35:4453–4455.

Lackey, A. C. R., and J. W. Boughman. 2017. Evolution of reproductive isolation in stickleback fish. Evolution. 71:357–372.

Lam, I., and S. Keeney. 2015. Nonparadoxical evolutionary stability of the recombination initiation landscape in yeast. Science. 350:932–937.

Lamichhaney, S., F. Han, M. T. Webster, L. Andersson, B. R. Grant, and P. R. Grant. 2018. Rapid hybrid speciation in Darwin's finches. Science. 359:224–228.

Langmead, B., and S. L. Salzberg. 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods 9:357.

Lawson, L. P., and K. Petren. 2017. The adaptive genomic landscape of beak morphology in Darwin's finches. Mol. Ecol. 26:4978–4989.

Leigh, J. W., and D. Bryant. 2015. POPART: full-feature software for haplotype network construction. Methods Ecol. Evol. 6:1110–1116.

Lenth, R., H. Singmann, and J. Love. 2018. Emmeans: Estimated marginal means, aka least-squares means. R Package version 1.

Li, H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv Prepr. arXiv1303.3997.

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

Li, J., B. Bed'hom, S. Marthey, M. Valade, A. Dureux, M. Moroldo, C. Péchoux, J.-L. Coville, D. Gourichon, A. Vieaud, B. Dorshorst, L. Andersson, and M. Tixier-Boichard. 2019. A missense mutation in TYRP1 causes the chocolate plumage color in chicken and alters melanosome structure. Pigment Cell Melanoma Res. 32:381–390.

Lichten, M., and A. S. H. Goldman. 1995. Meiotic recombination hotspots. Annu. Rev. Genet. 29:423–444.

Liedvogel, M., S. Åkesson, and S. Bensch. 2011. The genetics of migration on the move. Trends Ecol. Evol. 26:561–569.

Liedvogel, M., K. W. Larson, M. Lundberg, A. Gursoy, L. I. Wassenaar, K. A. Hobson, S. Bensch, and S. Åkesson. 2014. No evidence for assortative mating within a willow warbler migratory divide. Front. Zool. 11:52–61.

Liu, D., G. Zhang, H. Jiang, and J. Lu. 2018. Detours in long-distance migration across the Qinghai-Tibetan Plateau: individual consistency and habitat associations. PeerJ 6:e4304.

Lundberg, M., M. Liedvogel, K. Larson, H. Sigeman, M. Grahn, A. Wright, S. Åkesson, and S. Bensch. 2017. Genetic differences between willow warbler migratory phenotypes are few and cluster in large haplotype blocks. Evol. Lett. 1:155–168.

Lyons, J. I., A. A. Pierce, S. M. Barribeau, E. D. Sternberg, A. J. Mongue, and J. C. De Roode. 2012. Lack of genetic differentiation between monarch butterflies with divergent migration destinations. Mol. Ecol. 21:3433–3444.

Lyons, L. A., I. T. Foe, H. C. Rah, and R. A. Grahn. 2005. Chocolate coated cats: TYRP1 mutations for brown color in domestic cats. Mamm. Genome 16:356–366.

Ma, C., H. B. Vander Zanden, M. B. Wunder, and G. J. Bowen. 2020. assignR: An R package for isotope-based geographic assignment. Methods Ecol. Evol. 11:996–1001.

Maan, M. E., and O. Seehausen. 2011. Ecology, sexual selection and speciation. Ecol. Lett. 14:591–602.

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 Ecol. Evol. 4:906–913.

Marques, D. A., K. Lucek, J. I. Meier, S. Mwaiko, C. E. Wagner, L. Excoffier, and O. Seehausen. 2016. Genomics of Rapid Incipient Speciation in Sympatric Threespine Stickleback. PLOS Genet. 12:e1005887.

Marques, D. A., J. I. Meier, and O. Seehausen. 2019. A Combinatorial View on Speciation and Adaptive Radiation. Trends Ecol. Evol. 34:531–544.

Martin, S. H., K. K. Dasmahapatra, N. J. Nadeau, C. Salazar, J. R. Walters, F. Simpson, M. Blaxter, A. Manica, J. Mallet, and C. D. Jiggins. 2013. Genome-wide evidence for speciation with gene flow in *Heliconius* butterflies. Genome Res. 23:1817–1828.

Martin, S. H., J. W. Davey, C. Salazar, and C. D. Jiggins. 2019. Recombination rate variation shapes barriers to introgression across butterfly genomes. PLOS Biol. 17:e2006288.

Mason, N. A., and K. J. Burns. 2013. Molecular phylogenetics of the Neotropical seedeaters and seed-finches (*Sporophila*, *Oryzoborus*, *Dolospingus*). Ornitol. Neotrop. 24:139–155.

Matute, D. R. 2010. Reinforcement Can Overcome Gene Flow during Speciation in *Drosophila*. Curr. Biol. 20:2229–2233.

Matute, D. R., J. Gavin-Smyth, and G. Liu. 2014. Variable post-zygotic isolation in *Drosophila melanogaster*/*D. simulans* hybrids. J. Evol. Biol. 27:1691–1705.

Mayr, E. 1963. Animal species and evolution. Oxford University Press, London.

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. Mol. Ecol. 18:665–679.

McKellar, A. E., P. P. Marra, and L. M. Ratcliffe. 2013. Starting over: experimental effects of breeding delay on reproductive success in early-arriving male American redstarts. J. Avian Biol. 44:495–503.

McKenna, A., M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernytsky, K. Garimella, D. Altshuler, S. Gabriel, and M. Daly. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20:1297–1303.

Medolago, C. A. B., M. C. Costa, L. F. Silveira, and M. R. Francisco. 2020. Hybridization between two recently diverged Neotropical passerines: The Pearly-bellied Seedeater *Sporophila pileata*, and the Copper Seedeater *S. bouvreuil* (Aves, Passeriformes, Thraupidae). PLoS One 15:e0229714.

Meier, J. I., D. A. Marques, C. E. Wagner, L. Excoffier, and O. Seehausen. 2018. Genomics of Parallel Ecological Speciation in Lake Victoria Cichlids. Mol. Biol. Evol. 35:1489–1506.

Mendelson, T. C. 2003. Sexual isolation evolves faster than hybrid inviability in a diverse and sexually dimorphic genus of fish (Percidae: Etheostoma). Evolution. 57:317–327.

Mendelson, T. C., J. M. Gumm, M. D. Martin, and P. J. Ciccotto. 2017. Preference for conspecifics evolves earlier in males than females in a sexually dimorphic radiation of fishes. Evolution. 72:337–347.

Mendelson, T. C., V. E. Imhoff, and J. J. Venditti. 2007. The Accumulation of Reproductive Barriers During Speciation: Postmating Barriers in Two Behaviorally Isolated Species of Darters (Percidae: Etheostoma). Evolution. 61:2596–2606.

Meyer, M., T. Munzner, and H. Pfister. 2009. MizBee: a multiscale synteny browser. IEEE Trans. Vis. Comput. Graph. 15:897–904.

Miller-Butterworth, C. M., D. S. Jacobs, and E. H. Harley. 2003. Strong population substructure is correlated with morphology and ecology in a migratory bat. Nature 424:187–191.

Milner-Gulland, E. J., J. M. Fryxell, and A. R. E. Sinclair. 2011. Animal migration: a synthesis. Oxford University Press, New York.

Moore, A. J., P. A. Gowaty, W. G. Wallin, and P. J. Moore. 2001. Sexual conflict and the evolution of female mate choice and male social dominance. Proc. R. Soc. London. Ser. B Biol. Sci. 268:517–523.

Musiani, M., J. A. Leonard, H. D. Cluff, C. C. Gates, S. Mariani, P. C. Paquet, C. Vilà, and R. K. Wayne. 2007. Differentiation of tundra/taiga and boreal coniferous forest wolves: genetics, coat colour and association with migratory caribou. Mol. Ecol. 16:4149–4170.

Nachman, M. W., and B. A. Payseur. 2012. Recombination rate variation and speciation: theoretical predictions and empirical results from rabbits and mice. Philos. Trans. R. Soc. B Biol. Sci. 367:409–421.

Nosil, P. 2008. Speciation with gene flow could be common. Mol. Ecol. 17:2103–2106.

Nosil, P., S. P. Egan, and D. J. Funk. 2008. Heterogeneous genomic differentiation between walking-stick ecotypes: “Isolation by adaptation” and multiple roles for divergent selection. Evolution. 62:316–336.

Nosil, P., D. J. Funk, and D. Ortiz-Barrientos. 2009. Divergent selection and heterogeneous genomic divergence. Mol. Ecol. 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. Mol. Ecol. 6:955–970.

Okonechnikov, K., O. Golosova, M. Fursov, and the UGENE team. 2012. Unipro UGENE: a unified bioinformatics toolkit. Bioinformatics 28:1166–1167.

Oksanen, J., F. G. Blanchet, R. Kindt, P. Legendre, R. B. O'hara, G. L. Simpson, P. Solymos, M. H. H. Stevens, and H. Wagner. 2010. Vegan: community ecology package. R package version 1.17-4. URL http//CRAN. R-project. org/package= vegan.

Ortiz-Barrientos, D., J. Engelstädter, and L. H. Rieseberg. 2016. Recombination Rate Evolution and the Origin of Species. Trends Ecol. Evol. 31:226–236.

Panhuis, T. M., R. Butlin, M. Zuk, and T. Tregenza. 2001. Sexual selection and speciation. Trends Ecol. Evol. 16:364–371.

Patten, M. A., J. T. Rotenberry, and M. Zuk. 2004. Habitat selection, acoustic adaptation, and the evolution of reproductive isolation. Evolution. 58:2144–2155.

Peterson, B. K., J. N. Weber, E. H. Kay, H. S. Fisher, and H. E. Hoekstra. 2012. Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. PLoS One 7:e37135–e37135.

Pfeifer, B., U. Wittelsbürger, S. E. Ramos-Onsins, and M. J. Lercher. 2014. PopGenome: An Efficient Swiss Army Knife for Population Genomic Analyses in R. Mol. Biol. Evol. 31:1929–1936.

Pfennig, K. S., and A. M. Rice. 2014. Reinforcement generates reproductive isolation between neighbouring conspecific populations of spadefoot toads. Proc. R. Soc. B Biol. Sci. 281.

Pinho, C., and J. Hey. 2010. Divergence with Gene Flow: Models and Data. Annu. Rev. Ecol. Evol. Syst. 41:215–230.

Plummer, M., N. Best, K. Cowles, and K. Vines. 2006. CODA: convergence diagnosis and output analysis for MCMC. R news 6:7–11.

Poelstra, J. W., N. Vijay, C. M. Bossu, H. Lantz, B. Ryll, I. Müller, V. Baglione, P. Unneberg, M. Wikelski, and M. G. Grabherr. 2014. The genomic landscape underlying phenotypic integrity in the face of gene flow in crows. Science. 344:1410–1414.

Poelstra, J. W., N. Vijay, M. P. Hoeppner, and J. B. W. Wolf. 2015. Transcriptomics of colour patterning and coloration shifts in crows. Mol. Ecol. 24:4617–4628.

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

Prince, D. J., S. M. O'Rourke, T. Q. Thompson, O. A. Ali, H. S. Lyman, I. K. Saglam, T. J. Hotaling, A. P. Spidle, and M. R. Miller. 2017. The evolutionary basis of premature migration in Pacific salmon highlights the utility of genomics for informing conservation. Sci. Adv. 3:e1603198.

Pujolar, J. M., M. W. Jacobsen, T. D. Als, J. Frydenberg, E. Magnussen, B. Jónsson, X. Jiang, L. Cheng, D. Bekkevold, G. E. Maes, L. Bernatchez, and M. M. Hansen. 2014. Assessing patterns of hybridization between North Atlantic eels using diagnostic single-nucleotide polymorphisms. Heredity. 112:627–637.

Pulido-Santacruz, P., A. Aleixo, and J. T. Weir. 2018. Morphologically cryptic Amazonian bird species pairs exhibit strong postzygotic reproductive isolation. Proc. R. Soc. B Biol. Sci. 285:20172081.

Pulido, F., and P. Berthold. 2010. Current selection for lower migratory activity will drive the evolution of residency in a migratory bird population. Proc. Natl. Acad. Sci. 107:7341–7346.

Purcell, S., B. Neale, K. Todd-Brown, L. Thomas, M. A. R. Ferreira, D. Bender, J. Maller, P. Sklar, P. I. W. de Bakker, M. J. Daly, and P. C. Sham. 2007. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. Am. J. Hum. Genet. 81:559–575.

QGIS Development Team. 2019. QGIS Geographic Information System. Open Source Geospatial Foundation Project.

Quinn, T. P., M. J. Unwin, and M. T. Kinnison. 2000. Evolution of Temporal Isolation in the Wild: Genetic Divergence in Timing of Migration and Breeding by Introduced Chinook Salmon Populations. Evolution. 54:1372–1385.

Qvarnström, A., A. M. Rice, and H. Ellegren. 2010. Speciation in *Ficedula* flycatchers. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 365:1841–1852.

R Core Team. 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rakhimberdiev, E., A. Saveliev, T. Piersma, and J. Karagicheva. 2017. FLightR: an R package for reconstructing animal paths from solar geolocation loggers. Methods Ecol. Evol. 8:1482–1487.

Rakhimberdiev, E., N. R. Senner, M. A. Verhoeven, D. W. Winkler, W. Bouten, and T. Piersma. 2016. Comparing inferences of solar geolocation data against high-precision GPS data: annual movements of a double-tagged black-tailed godwit. J. Avian Biol. 47:589–596.

Randler, C. 2002. Avian hybridization, mixed pairing and female choice. Anim. Behav. 63:103–119.

Rappole, J. H., and A. R. Tipton. 1991. New Harness Design for Attachment of Radio Transmitters to Small Passerines (Nuevo Diseño de Arnés para Atar Transmisores a

Passeriformes Pequeños). J. F. Ornithol. 62:335–337.

Ravinet, M., R. Faria, R. K. Butlin, J. Galindo, N. Bierne, M. Rafajlović, M. A. F. Noor, B. Mehlig, and A. M. Westram. 2017. Interpreting the genomic landscape of speciation: a road map for finding barriers to gene flow. J. Evol. Biol. 30:1450–1477.

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. Nat. Commun. 4:1827.

Repenning, M., and C. S. Fontana. 2013. A New Species of Gray Seedeater (Emberizidae: *Sporophila*) from Upland Grasslands of Southern Brazil. Auk 130:791–803.

Revell, L. J. 2012. phytools: an R package for phylogenetic comparative biology (and other things). Methods Ecol. Evol. 3:217–223.

Rieseberg, L. H. 2001. Chromosomal rearrangements and speciation. Trends Ecol. Evol. 16:351–358.

Ritchie, M. G. 2007. Sexual Selection and Speciation. Annu. Rev. Ecol. Evol. Syst. 38:79–102.

Rochette, N. C., A. G. Rivera-Colón, and J. M. Catchen. 2019. Stacks 2: Analytical methods for paired-end sequencing improve RADseq-based population genomics. Mol. Ecol. 28:4737–4754.

Roesti, M., A. P. Hendry, W. Salzburger, and D. Berner. 2012. Genome divergence during evolutionary diversification as revealed in replicate lake–stream stickleback population pairs. Mol. Ecol. 21:2852–2862.

Rohland, N., and D. Reich. 2012. Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. Genome Res. 22:939–946.

Rohwer, S., and D. E. Irwin. 2011. Molt, Orientation, and Avian Speciation. Auk Ornithol. Adv.

128:419–425.

Rolshausen, G., K. A. Hobson, and H. M. Schaefer. 2010. Spring arrival along a migratory divide of sympatric blackcaps (*Sylvia atricapilla*). Oecologia 162:175.

Rolshausen, G., G. Segelbacher, C. Hermes, K. A. Hobson, and H. M. Schaefer. 2013. Individual differences in migratory behavior shape population genetic structure and microhabitat choice in sympatric blackcaps (*Sylvia atricapilla*). Ecol. Evol. 3:4278–4289.

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. Curr. Biol. 19:2097–2101.

Rosenblum, E. B., B. A. J. Sarver, J. W. Brown, S. Des Roches, K. M. Hardwick, T. D. Hether, J. M. Eastman, M. W. Pennell, and L. J. Harmon. 2012. Goldilocks Meets Santa Rosalia: An Ephemeral Speciation Model Explains Patterns of Diversification Across Time Scales. Evol. Biol. 39:255–261.

Rubenstein, D. R., and K. A. Hobson. 2004. From birds to butterflies: animal movement patterns and stable isotopes. Trends Ecol. Evol. 19:256–263.

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, and H. Slabbekoorn. 2012. Differences in timing of migration and response to sexual signalling drive asymmetric hybridization across a migratory divide. J. Evol. Biol. 25:1741–1750.

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*). Proc. R. Soc. London. Ser. B Biol. Sci. 269:1375–81.

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

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

Ryan, G., D. Shannon, A. Lenin, and T. Michael. 2016. Adaptive, but not condition-dependent, body shape differences contribute to assortative mating preferences during ecological speciation. Evolution. 70:2809–2822.

Safran, R. J., E. S. C. Scordato, L. B. Symes, R. L. Rodríguez, and T. C. Mendelson. 2013. Contributions of natural and sexual selection to the evolution of premating reproductive isolation: a research agenda. Trends Ecol. Evol. 28:643–650.

Safran, R. J., E. S. C. Scordato, M. R. Wilkins, J. K. Hubbard, B. R. Jenkins, T. Albrecht, S. M. Flaxman, H. Karaardıç, Y. Vortman, A. Lotem, P. Nosil, P. Pap, S. Shen, S.-F. Chan, T. L. Parchman, and N. C. Kane. 2016. Genome-wide differentiation in closely related populations: the roles of selection and geographic isolation. Mol. Ecol. 25:3865–3883.

Saino, N., T. Szép, R. Ambrosini, M. Romano, and A. P. Møller. 2004. Ecological conditions during winter affect sexual selection and breeding in a migratory bird. Proc. R. Soc. London. Ser. B Biol. Sci. 271:681–686.

Samuk, K., G. L. Owens, K. E. Delmore, S. E. Miller, D. J. Rennison, and D. Schluter. 2017. Gene flow and selection interact to promote adaptive divergence in regions of low recombination. Mol. Ecol. 26:4378–4390.

Sánchez-Guillén, R. A., M. Wellenreuther, and A. Cordero Rivera. 2012. Strong Asymmetry in the Relative Strengths of Prezygotic and Postzygotic Barriers Between Two Damselfly Sister Species. Evolution. 66:690–707.

Sardell, J. M., and M. Kirkpatrick. 2020. Sex Differences in the Recombination Landscape. Am. Nat. 195:361–379.

Schluter, D. 2009. Evidence for Ecological Speciation and Its Alternative. Science. 323:737–741.

Schubert, M., S. Lindgreen, and L. Orlando. 2016. AdapterRemoval v2: rapid adapter trimming, identification, and read merging. BMC Res. Notes 9:88.

Schumer, M., D. L. Powell, P. J. Delclós, M. Squire, R. Cui, P. Andolfatto, and G. G. Rosenthal. 2017. Assortative mating and persistent reproductive isolation in hybrids. Proc. Natl. Acad. Sci. 114:10936–10941.

Scordato, E. S. C., and R. J. Safran. 2014. Geographic variation in sexual selection and implications for speciation in the Barn Swallow. Avian Res. 5:8.

Scordato, E. S. C., C. C. R. Smith, G. A. Semenov, Y. Liu, M. R. Wilkins, W. Liang, A. Rubtsov, G. Sundev, K. Koyama, S. P. Turbek, M. B. Wunder, C. A. Stricker, and R. J. Safran. 2020. Migratory divides coincide with reproductive barriers across replicated avian hybrid zones above the Tibetan Plateau. Ecol. Lett. 23:231–241.

Scordato, E. S. C., M. R. Wilkins, G. Semenov, A. S. Rubtsov, N. C. Kane, and R. J. Safran. 2017. Genomic variation across two barn swallow hybrid zones reveals traits associated with divergence in sympatry and allopatry. Mol. Ecol. 26:5676–5691.

Secondi, J., M. Okassa, S. Sourice, and M. Théry. 2014. Habitat-dependent species recognition in hybridizing newts. Evol. Biol. 41:71–80.

Seehausen, O. 2004. Hybridization and adaptive radiation. Trends Ecol. Evol. 19:198–207.

Seehausen, O., R. K. Butlin, I. Keller, C. E. Wagner, J. W. Boughman, P. A. Hohenlohe, C. L. Peichel, G.-P. Saetre, C. Bank, Å. Brännström, A. Brelsford, C. S. Clarkson, F.

Eroukhmanoff, J. L. Feder, M. C. Fischer, A. D. Foote, P. Franchini, C. D. Jiggins, F. C. Jones, A. K. Lindholm, K. Lucek, M. E. Maan, D. A. Marques, S. H. Martin, B. Matthews, J. I. Meier, M. Möst, M. W. Nachman, E. Nonaka, D. J. Rennison, J. Schwarzer, E. T. Watson, A. M. Westram, and A. Widmer. 2014. Genomics and the origin of species. Nat. Rev. Genet. 15:176.

Seehausen, O., Y. Terai, I. S. Magalhaes, K. L. Carleton, H. D. J. Mrosso, R. Miyagi, I. van der Sluijs, M. V Schneider, M. E. Maan, H. Tachida, H. Imai, and N. Okada. 2008. Speciation through sensory drive in cichlid fish. Nature 455:620–626.

Seehausen, O., J. J. M. van Alphen, and F. Witte. 1997. Cichlid fish diversity threatened by eutrophication that curbs sexual selection. Science. 277:1808–1811.

Selz, O. M., M. E. R. Pierotti, M. E. Maan, C. Schmid, and O. Seehausen. 2014. Female preference for male color is necessary and sufficient for assortative mating in 2 cichlid sister species. Behav. Ecol. 25:612–626.

Semenov, G. A., E. A. Koblik, Y. A. Red'kin, and A. V Badyaev. 2018. Extensive phenotypic diversification coexists with little genetic divergence and a lack of population structure in the White Wagtail subspecies complex (*Motacilla alba*). J. Evol. Biol. 31:1093–1108.

Semenov, G. A., E. Linck, E. D. Enbody, R. B. Harris, D. R. Khaydarov, P. Alström, L. Andersson, and S. A. Taylor. 2021. Asymmetric introgression reveals the genetic architecture of a plumage trait. Nat. Commun. 12:1019.

Semenov, G. A., R. J. Safran, C. C. R. Smith, S. P. Turbek, S. P. Mullen, and S. M. Flaxman. 2019. Unifying Theoretical and Empirical Perspectives on Genomic Differentiation. Trends Ecol. Evol. 34:987–995.

Semenov, G. A., E. S. C. Scordato, D. R. Khaydarov, C. C. R. Smith, N. C. Kane, and R. J.

Safran. 2017. Effects of assortative mate choice on the genomic and morphological structure of a hybrid zone between two bird subspecies. Mol. Ecol. 26:6430–6444.

Servedio, M. R., and J. W. Boughman. 2017. The Role of Sexual Selection in Local Adaptation and Speciation. Annu. Rev. Ecol. Evol. Syst. 48:85–109.

Shafer, A. B. A., J. M. Northrup, M. Wikelski, G. Wittemyer, and J. B. W. Wolf. 2016. Forecasting Ecological Genomics: High-Tech Animal Instrumentation Meets High-Throughput Sequencing. PLOS Biol. 14:e1002350.

Shanfelter, A. F., S. L. Archambeault, and M. A. White. 2019. Divergent fine-scale recombination landscapes between a freshwater and marine population of threespine stickleback fish. Genome Biol. Evol. 11:1552.

Short, L. L. 1971. Aves nuevas o poco comunes de Corrientes. Rev. Mus. Argent. Cienc. Nat."Bernardino Rivadavia" Inst. Nac. Invest. Cienc. Nat. Zool 9:283–309.

Siepel, A., G. Bejerano, J. S. Pedersen, A. S. Hinrichs, M. Hou, K. Rosenbloom, H. Clawson, J. Spieth, L. W. Hillier, S. Richards, G. M. Weinstock, R. K. Wilson, R. A. Gibbs, W. J. Kent, W. Miller, and D. Haussler. 2005. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. Genome Res. 15:1034–1050.

Singhal, S., E. M. Leffler, K. Sannareddy, I. Turner, O. Venn, D. M. Hooper, A. I. Strand, Q. Li, B. Raney, C. N. Balakrishnan, S. C. Griffith, G. McVean, and M. Przeworski. 2015. Stable recombination hotspots in birds. Science. 350:928–932.

Smagulova, F., K. Brick, Y. Pu, R. D. Camerini-Otero, and G. V Petukhova. 2016. The evolutionary turnover of recombination hot spots contributes to speciation in mice. Genes Dev. 30:266–280.

Smeds, L., A. Qvarnström, and H. Ellegren. 2016. Direct estimate of the rate of germline

mutation in a bird. Genome Res. 26:1211–1218.

Smith, C. C. R., S. M. Flaxman, E. S. C. Scordato, N. C. Kane, A. K. Hund, B. M. Sheta, and R. J. Safran. 2018. Demographic inference in barn swallows using whole-genome data shows signal for bottleneck and subspecies differentiation during the Holocene. Mol. Ecol. 27:4200–4212.

Smukowski, C. S., and M. A. F. Noor. 2011. Recombination rate variation in closely related species. Heredity. 107:496–508.

Sobel, J. M., and M. A. Streisfeld. 2015. Strong premating reproductive isolation drives incipient speciation in *Mimulus aurantiacus*. Evolution. 69:447–461.

Stamatakis, A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30:1312–1313.

Stapley, J., P. G. D. Feulner, S. E. Johnston, A. W. Santure, and C. M. Smadja. 2017. Variation in recombination frequency and distribution across eukaryotes: patterns and processes. Philos. Trans. R. Soc. B Biol. Sci. 372:20160455.

Stephan, W., L. Xing, D. A. Kirby, and J. M. Braverman. 1998. A test of the background selection hypothesis based on nucleotide data from Drosophila ananassae. Proc. Natl. Acad. Sci. 95:5649–5654.

Still, C. J., J. A. Berry, G. J. Collatz, and R. S. DeFries. 2003. Global distribution of C3 and C4 vegetation: Carbon cycle implications. Global Biogeochem. Cycles 17:6–14.

Strimas-Mackey, M., E. Miller, and W. Hochachka. 2017. auk: eBird Data Extraction and Processing with AWK. R Package version 0.03.

Stryjewski, K. F., and M. D. Sorenson. 2017. Mosaic genome evolution in a recent and rapid avian radiation. Nat. Ecol. Evol. 1:1912.

Stukenbrock, E. H., and J. Y. Dutheil. 2018. Fine-Scale Recombination Maps of Fungal Plant Pathogens Reveal Dynamic Recombination Landscapes and Intragenic Hotspots. Genetics 208:1209–1229.

Sturm, R. A., and T. N. Frudakis. 2004. Eye colour: portals into pigmentation genes and ancestry. Trends Genet. 20:327–332.

Sullivan, B. L., J. L. Aycrigg, J. H. Barry, R. E. Bonney, N. Bruns, C. B. Cooper, T. Damoulas, A. A. Dhondt, T. Dietterich, A. Farnsworth, D. Fink, J. W. Fitzpatrick, T. Fredericks, J. Gerbracht, C. Gomes, W. M. Hochachka, M. J. Iliff, C. Lagoze, F. A. La Sorte, M. Merrifield, W. Morris, T. B. Phillips, M. Reynolds, A. D. Rodewald, K. V Rosenberg, N. M. Trautmann, A. Wiggins, D. W. Winkler, W.-K. Wong, C. L. Wood, J. Yu, and S. Kelling. 2014. The eBird enterprise: An integrated approach to development and application of citizen science. Biol. Conserv. 169:31–40.

Svedin, N., C. Wiley, T. Veen, L. Gustafsson, and A. Qvarnström. 2008. Natural and sexual selection against hybrid flycatchers. Proc. R. Soc. B Biol. Sci. 275:735–744.

Taborsky, B., L. Guyer, and P. Demus. 2014. 'Prudent habitat choice': a novel mechanism of size-assortative mating. J. Evol. Biol. 27:1217–1228.

Tavares, H. 2020. windowscanr: Apply functions using sliding windows. R package version 0.1.

Taylor, R. S., and V. L. Friesen. 2017. The role of allochrony in speciation. Mol. Ecol. 26:3330–3342.

Thrasher, D. J., B. G. Butcher, L. Campagna, M. S. Webster, and I. J. Lovette. 2018. Double-digest RAD sequencing outperforms microsatellite loci at assigning paternity and estimating relatedness: A proof of concept in a highly promiscuous bird. Mol. Ecol. Resour. 18:953–965.

Tinghitella, R. M., A. C. R. Lackey, M. Martin, P. D. Dijkstra, J. P. Drury, R. Heathcote, J. Keagy, E. S. C. Scordato, and A. M. Tyers. 2017. On the role of male competition in speciation: a review and research agenda. Behav. Ecol. 29:783–797.

Toews, D. P. L., S. A. Taylor, R. Vallender, A. Brelsford, B. G. Butcher, P. W. Messer, and I. J. Lovette. 2016. Plumage genes and little else distinguish the genomes of hybridizing warblers. Curr. Biol. 26:2313–2318.

Tonnis, B., P. R. Grant, B. R. Grant, and K. Petren. 2005. Habitat selection and ecological speciation in Galápagos warbler finches (*Certhidea olivacea* and *Certhidea fusca*). Proc. R. Soc. B Biol. Sci. 272:819–826.

Turbek, S. P., M. Browne, C. Pasian, and A. S. Di Giacomo. 2019. First nest description of the Iberá Seedeater (*Sporophila iberaensis*). Wilson J. Ornithol. 131:156–160.

Turbek, S. P., E. S. C. Scordato, and R. J. Safran. 2018. The Role of Seasonal Migration in Population Divergence and Reproductive Isolation. Trends Ecol. Evol. 33:164–175.

Turbek, S. P., M. Browne, A. S. Di Giacomo, C. Kopuchian, W. M. Hochachka, C. Estalles, D. A. Lijtmaer, P. L. Tubaro, L. F. Silveira, I. J. Lovette, R. J. Safran, S. A. Taylor, and L. Campagna. 2021. Science 371:eabc0256.

Turner, A. 2010. The barn swallow. T & A D Poyser, London.

Turner, S. 2018. qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. J. Open Source Softw. 3:731.

Turner, T. L., M. W. Hahn, and S. V Nuzhdin. 2005. Genomic Islands of Speciation in *Anopheles gambiae*. PLOS Biol. 3:e285.

Uy, J. A. C., D. E. Irwin, and M. S. Webster. 2018. Behavioral Isolation and Incipient Speciation in Birds. Annu. Rev. Ecol. Evol. Syst. 49:1–24.

Uy, J. A. C., R. G. Moyle, and C. E. Filardi. 2009a. Plumage and Song Differences Mediate Species Recognition Between Incipient Flycatcher Species of the Solomon Islands. Evolution. 63:153–164.

Uy, J. A. C., R. G. Moyle, C. E. Filardi, and Z. A. Cheviron. 2009b. Difference in plumage color used in species recognition between incipient species is linked to a single amino acid substitution in the melanocortin-1 receptor. Am. Nat. 174:244–254.

Verzijden, M. N., C. ten Cate, M. R. Servedio, G. M. Kozak, J. W. Boughman, and E. I. Svensson. 2012. The impact of learning on sexual selection and speciation. Trends Ecol. Evol. 27:511–519.

Visser, M., M. Kayser, F. Grosveld, and R.-J. Palstra. 2014. Genetic variation in regulatory DNA elements: the case of OCA2 transcriptional regulation. Pigment Cell Melanoma Res. 27:169–177.

Visser, M., M. Kayser, and R.-J. Palstra. 2012. HERC2 rs12913832 modulates human pigmentation by attenuating chromatin-loop formation between a long-range enhancer and the OCA2 promoter. Genome Res. 22:446–455.

Wallbank, R. W. R., S. W. Baxter, C. Pardo-Diaz, J. J. Hanly, S. H. Martin, J. Mallet, K. K. Dasmahapatra, C. Salazar, M. Joron, N. Nadeau, W. O. McMillan, and C. D. Jiggins. 2016. Evolutionary Novelty in a Butterfly Wing Pattern through Enhancer Shuffling. PLOS Biol. 14:e1002353.

Wang, I. J. 2013. Examining the full effects of landscape heterogeneity on spatial genetic variation: a multiple matrix regression approach for quantifying geographic and ecological isolation. Evolution. 67:3403–3411.

Wassenaar, L. I., and K. A. Hobson. 2003. Comparative equilibration and online technique for

determination of non-exchangeable hydrogen of keratins for use in animal migration studies. Isotopes Environ. Health Stud. 39:211–217.

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

West-Eberhard, M. J. 1983. Sexual Selection, Social Competition, and Speciation. Q. Rev. Biol. 58:155–183.

Wheatcroft, D., and A. Qvarnström. 2017. Genetic divergence of early song discrimination between two young songbird species. Nat. Ecol. Evol. 1:192.

Williams, T. H., and T. C. Mendelson. 2014. Quantifying reproductive barriers in a sympatric pair of darter species. Evol. Biol. 41:212–220.

Winker, K. 2010. On the origin of species through heteropatric differentiation: a review and a model of speciation in migratory animals. Ornithol. Monogr. 69:1–30.

Wood, C. C., J. W. Bickham, R. John Nelson, C. J. Foote, and J. C. Patton. 2008. Recurrent evolution of life history ecotypes in sockeye salmon: implications for conservation and future evolution. Evol. Appl. 1:207–221.

Wood, S. N. 2011. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. J. R. Stat. Soc. Ser. B. 73:3–36.

Wotherspoon, S., M. Sumner, and S. Lisovski. 2016. TwGeos: Basic data processing for light-level geolocation archival tags. Version 0.0-1.

Wunder, M. B. 2010. Using Isoscapes to Model Probability Surfaces for Determining Geographic Origins. Pp. 251–270 *in* J. B. West, G. J. Bowen, T. E. Dawson, and K. P. Tu, eds. Isoscapes: Understanding movement, pattern, and process on Earth through isotope mapping. Springer Netherlands, Dordrecht.

Xu, M., and K. L. Shaw. 2019. Genetic coupling of signal and preference facilitates sexual isolation during rapid speciation. Proc. R. Soc. B Biol. Sci. 286:20191607.

Xue, H.-J., W.-Z. Li, and X.-K. Yang. 2014. Assortative mating between two sympatric closely-related specialists: inferred from molecular phylogenetic analysis and behavioral data. Sci. Rep. 4:5436.

Zheng, X., D. Levine, J. Shen, S. M. Gogarten, C. Laurie, and B. S. Weir. 2012. A high-performance computing toolset for relatedness and principal component analysis of SNP data. Bioinformatics 28:3326–3328.

APPENDIX A3

Supplemental Figures and Tables for Chapter 3

SUPPLEMENTAL FIGURES

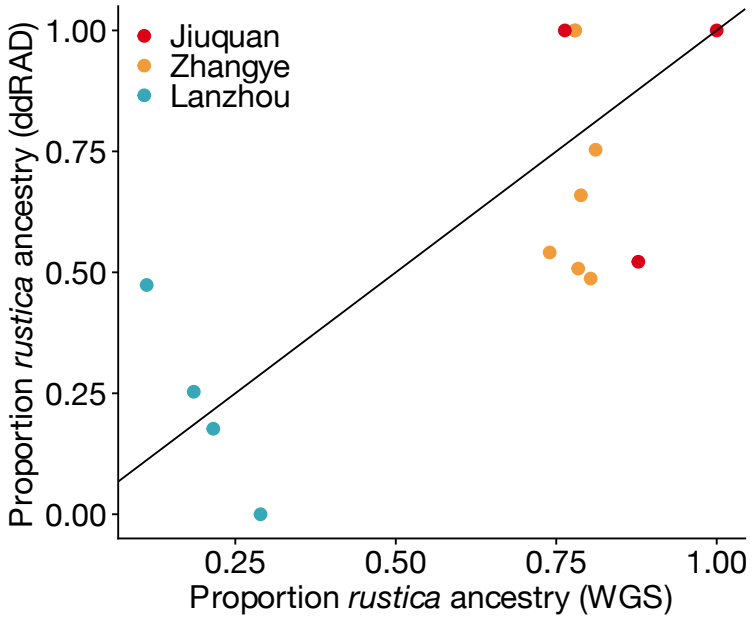


Figure S3.1. Comparison of WGS and ddRAD-based estimates of *H. r. rustica* ancestry for barn swallows tracked using geolocators in Gansu, China (one device failed to record data). The line indicates a 1:1 relationship between the two variables. We found a significant positive correlation between the proportion of *H. r. rustica* ancestry inferred using the two datasets (Pearson's r = 0.718, p = 0.004, n = 14).

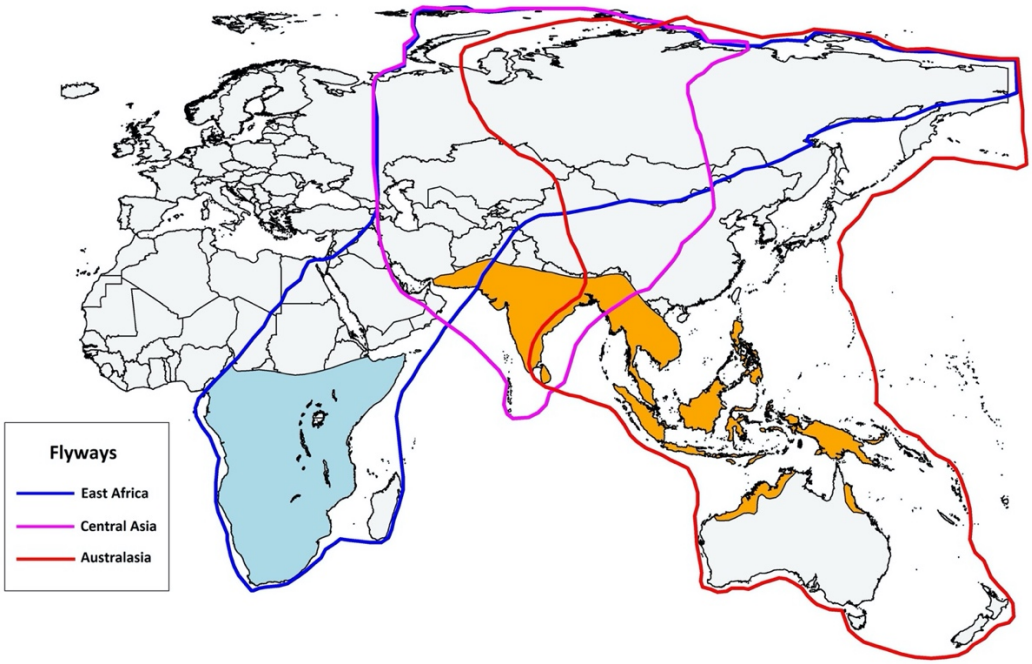


Figure S3.2. Map showing the migratory flyways used to restrict the possible wintering locations in the assignment model. Individuals were assigned a probability of overwintering in Africa (light blue) vs. Asia (orange).

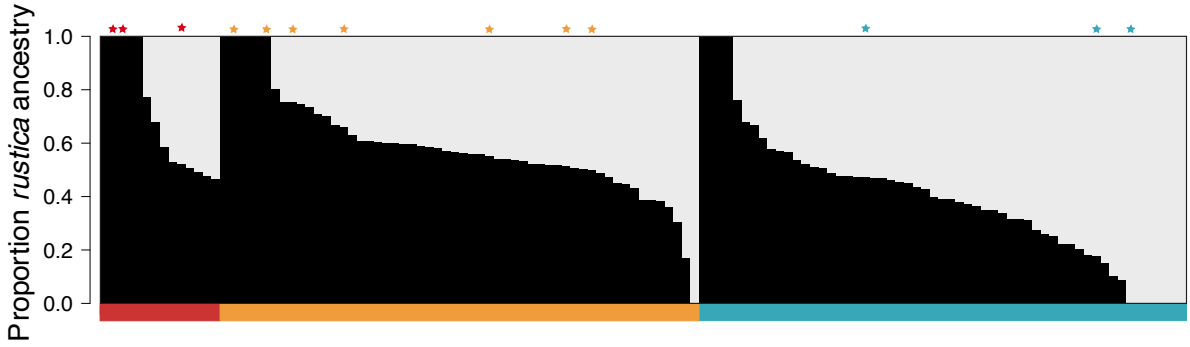


Figure S3.3. Admixture proportions generated using ddRAD data for two subspecies of barn swallow (*H. r. rustica* and *H. r. gutturalis*) breeding on either end of a hybrid zone in China (n = 127). The colors along the x axis indicate sampling location and correspond to cities on the map in Figure 3.1. The admixture proportions of the 13 individuals with geolocator tracks are indicated with stars.

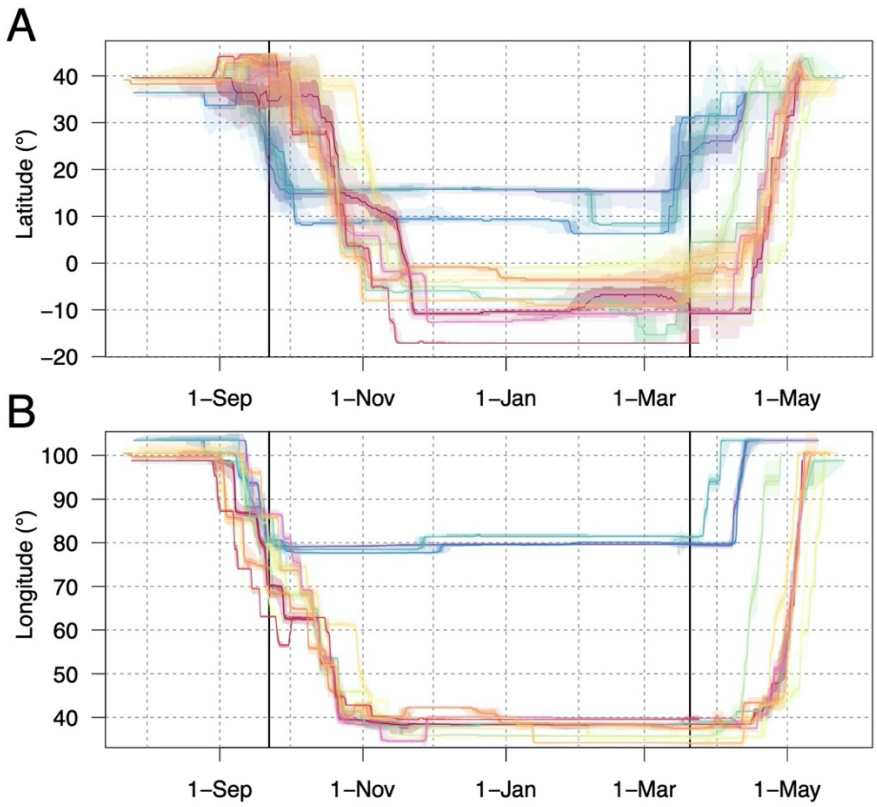


Figure S3.4. (**A**) Latitudinal and (**B**) longitudinal estimates throughout the annual cycle for 13 barn swallows tracked using geolocators. Individuals that traveled to Africa are shown in warm colors and birds that migrated to India are shown in cool colors. Median positions are represented as solid lines and 95% and 50% credible intervals are depicted as medium- and light-colored polygons, respectively, surrounding the flight paths. The two vertical black lines indicate the equinoxes.

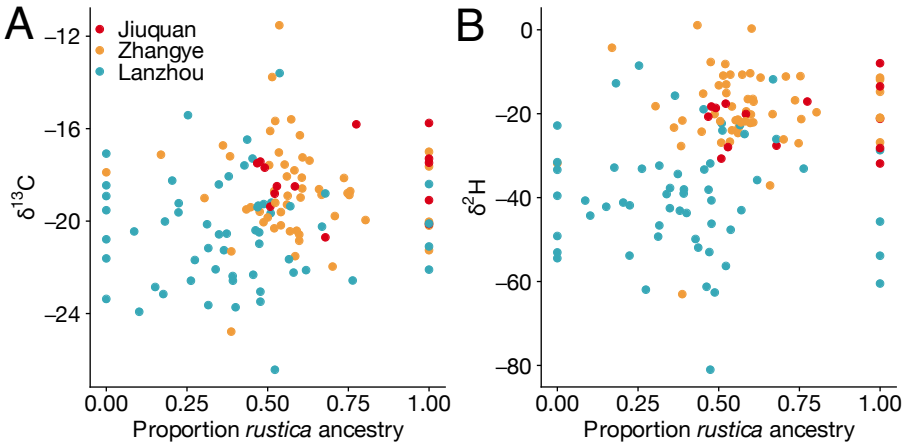


Figure S3.5. The relationship between proportion of *H. r. rustica* ancestry derived from ddRAD data and the feather **(A)** stable-carbon isotope ($\delta^{13}C$) and **(B)** stable-hydrogen isotope ($\delta^{2}H$) values of barn swallows (*H. r. rustica* and *H. r. gutturalis*) breeding across a hybrid zone in

China (n = 126). Colors correspond to cities on the map in Figure 3.1. We found a marginally significant positive correlation between proportion of *H. r. rustica* ancestry and $\delta^{13}C$ value (r = 0.152, p = 0.089, n = 126) and a significant positive correlation between proportion of *H. r. rustica* ancestry and $\delta^{2}H$ value (r = 0.290, p = 0.001, n = 126).

SUPPLEMENTAL TABLES

Table S3.1. Sample locations for the 168 individuals for which we generated WGS data. While all individuals were included in the dataset during variant calling to improve genotyping accuracy, we restricted the final dataset to individuals of *H. r. rustica and H r. gutturalis* that were sampled in Asia (Russia, China, Mongolia, and Japan) and excluded an additional region of admixture between *H. r. rustica* and *H r. gutturalis* in Russia. The restricted dataset (n = 78) was used to calculate admixture proportions for the 13 birds with geolocator tracks.

Country	Location	Latitude	Longitude	n
China	Urumqi	43.81749	87.62793	2
China	Dunhuang	40.09302	94.67034	1
China	Yumen	40.28390	97.03056	1
China	Jiuquan	39.71938	98.50991	3
China	Gaotai	39.36141	99.81342	1
China	Zhangye	38.98859	100.43651	7
China	Wuwei	37.88896	102.62696	5
China	Lanzhou	36.23648	103.43309	4
China	Yinchaun	38.52959	106.17479	1
China	Nanning	22.79192	108.32630	2
China	Xian	34.34565	108.80447	2
China	Hainan	19.22640	109.14651	3
China	Baotu	40.55583	110.00383	2
China	Changsha	28.40072	112.80742	3
China	Zhengzhou	34.82433	113.67071	2
China	Beijing	39.81956	116.33291	1
China	Qinhuangdao	39.92558	119.59475	2
China	Shenyang	41.80877	123.52697	1
China	Qiqihar	47.34004	123.97589	2
China	Changchun	43.89489	125.32068	2
China	Harbin	45.76124	126.60906	2
China	Shuangyashan	46.65299	131.14805	2

Egypt	Damietta	31.41366	31.56740	8
Israel	Amiad	32.92820	35.54070	4
Israel	Hoquq	32.92820	35.54070	2
Israel	Kahal	32.92820	35.54070	2
Japan	Tokyo	35.95223	139.12584	2
Japan	Hokkaido	42.30963	142.52504	3
Mongolia	Khovd	48.00975	91.66156	2
Mongolia	Durgun	48.33142	92.63147	3
Mongolia	Ugii Lake	47.75090	102.77232	2
Mongolia	Kharkhonh	47.19171	102.85620	1
Mongolia	Dashinchilin	47.84553	104.05680	3
Mongolia	Lun	47.86650	105.21169	2
Mongolia	Zuunmod	47.69861	106.95476	1
Mongolia	Erdene	47.71629	107.80251	1
Mongolia	Cincer Mandel Som	47.69396	108.46947	2
Mongolia	Tsenhermandal	47.65935	109.13263	1
Mongolia	Chingis Khan	47.33190	110.68623	2
Mongolia	Bayan Uul	49.12203	112.67764	2
Mongolia	Norovlin	48.46979	113.11306	1
Mongolia	Bulgan Soum Dornod	48.00061	113.93633	2
Mongolia	Dashbalbar	49.50925	114.45797	2
Morocco	Agadir	30.05334	-9.65147	2
Morocco	Marrakech	31.72004	-7.74009	2
Morocco	Beni-Mellal	32.32765	-6.50300	2
Morocco	Moulay	34.81042	-6.30846	1
Morocco	Souk	34.68093	-5.99680	1
Russia	Moscow	56.76593	37.78463	4
Russia	Yekaterinburg	57.54073	62.71657	3
Russia	Karasuk	53.93488	77.73870	2
Russia	Krasny Yar	55.72532	86.15158	1
Russia	Krasnoyarsk	56.35105	93.05357	2
Russia	Byronovka	55.88492	98.13112	1
Russia	Alzamay	55.59150	98.60941	2
Russia	Zamzor	55.37321	98.65220	4
Russia	Mara	55.00064	98.84569	1

Russia	Kamenka	55.00048	98.84857	2
Russia	Uk	55.07751	98.87161	2
Russia	Hingui Station	54.79531	99.34366	1
Russia	Hingui Village	54.79843	99.44060	3
Russia	Male-Kutulyk	53.40102	102.78564	2
Russia	Malamolevo	53.39322	102.86215	2
Russia	Zakaltoose	52.02126	106.59094	6
Russia	Nikolaevska	51.06395	111.79121	1
Russia	Tataurova	51.60738	112.93938	2
Russia	Mixed Barns	50.46408	113.45374	2
Russia	Narin Talacha	51.93720	114.96832	3
Russia	Civaky	52.83970	126.58935	1
Russia	Vozhaevka	50.74161	128.72934	2
Russia	Yadrina	48.96948	131.01845	1
Russia	Chernaevka	44.33296	132.51778	2
Russia	Magelevka	47.96813	134.91274	1
USA	Colorado	40.02740	-105.25190	8

Table S3.2. Sample locations for the 127 individuals included in the ddRAD dataset after filtering steps, which included the 13 individuals with geolocator tracks. To call variants, we combined this dataset with eight individuals from the broader WGS dataset that had nearly 100% assignment to *H. r. rustica* or *H. r. gutturalis*.

Location	**Latitude**	**Longitude**	**n**
Jiuquan	39.71938	98.50991	14
Zhangye	38.98859	100.43651	56
Lanzhou	36.23648	103.43309	57

APPENDIX A4

Supplemental Methods, Text, Figures, and Tables for Chapter 4

SUPPLEMENTAL METHODS

Field methods

Research was approved by the University of Colorado Institutional Animal Care and Use Committee (IACUC) protocols no. 2498 and 2683 and conducted in accordance with USFWS permit MB12129A-10, USDA permit 112702, and permission from the Ministerio de Ambiente y Desarrollo Sostenible de la Nación (Argentina), the Administración de Parques Nacionales (Argentina), and the Dirección de Parques y Reservas (Province of Corrientes, Argentina).

Whole-genome resequencing and variant discovery

Individually barcoded libraries were prepared with 200 ng of each sample and an insert size of 550 bp following the TruSeq Nano DNA library preparation kit protocol (Illumina). The 37 libraries were then pooled into two groups of 19 and 18 samples using concentrations of adapter-ligated DNA determined through digital polymerase chain reaction (PCR). We sequenced the pooled libraries on two Illumina NextSeq 500 lanes at the Cornell University Biotechnology Resource Center (BRC). Whole-genome resequencing generated over 860 million paired-end reads with a length of 151 bp, producing an expected per-individual coverage ranging between 3.9× and 10.7× (median: 5.4×; Table S4.2).

Double-digest restriction site-associated DNA sequencing

We digested 100-500 ng of DNA from each individual with SbfI and MspI (NEB) and ligated one of 20 P1 adapters (containing unique inline barcodes) and a P2-MspI adapter to each sample.

The samples were pooled together in groups of 20, purified with 1.5× volumes of magnetic beads made with Sera-Mag Magnetic Speed-beads (FisherSci), as detailed in (Rohland and Reich 2012), and size-selected for 450-600 bp fragments using BluePippin (Sage Science) at the Cornell University BRC. We added the Illumina index group sequences and sequencing adapters following size selection by performing 11 PCR cycles with Phusion DNA Polymerase (NEB). The products of the reactions were purified with 0.7 volumes of magnetic beads, and pooled in equimolar ratios to generate two libraries for sequencing. Each library was sequenced on one lane of an Illumina HiSeq 2500 at the BRC to generate 101-bp single-end reads.

Demographic modeling

We generated demographic inferences using G-PhoCS (Generalized Phylogenetic Coalescent Sampler) version 1.3 (Gronau et al. 2011), which allowed us to co-estimate effective population sizes, splitting times, and bi-directional migration rates (*i.e.,* introgression). We decided to conduct this analysis using ddRAD data, which provides two phased haplotypes per individual for a sub-sample of loci across the genome. This strategy allowed us to combine the ddRAD data generated for *S. iberaensis* and *S. hypoxantha* in this study ($N = 206$ individuals), with previously generated ddRAD data for 70 individuals belonging to six southern capuchino species from (Campagna et al. 2015) (*S. bouvreuil*, *S. hypoxantha*, *S. melanogaster*, *S. pileata*, *S. ruficollis*, and *S. palustris*). We first trimmed the 3' end of the sequences from (Campagna et al. 2015) to a length of 90 bp (from an original 132 bp after filtering and demultiplexing) to match the locus size generated in this study for *S. iberaensis* and *S. hypoxantha*. We subsequently aligned these data to the *S. hypoxantha* reference genome and assembled ddRAD loci in Stacks as described in the *Double-digest restriction site-associated DNA sequencing* section. We

discarded individuals with more than 10% missing data and exported loci present in at least 90% of all individuals, which allowed us to retain a total of 5,852 variant and invariant (*i.e.,* without SNPs) loci for 256 individuals, containing a total of 54,698 SNPs. This reference-based assembly allowed us to recover ~2,000 more loci than what was previously used for demographic modeling by Campagna *et al.* (Campagna et al. 2015). We explored the relationships among individuals and species, and compared the patterns to those observed from whole-genome sequencing data, by performing PCAs in *SNPRelate*. Prior to conducting the G-PhoCS analysis, we removed the invariant enzyme cut site from the 5' end of all loci.

The G-PhoCS analysis is based on the phylogenetic relationships among the species included in the demographic reconstruction. Because these relationships cannot be fully resolved within the rapid radiation of southern capuchinos (*e.g.,* see Figure 4.6A), we devised a strategy that compared species by pairs (see Figure 6 in Campagna *et al.* (2015)). We conducted a total of 16 G-PhoCS runs, including all pairwise combinations of six southern capuchino species and *S. bouvreuil* as an outgroup. We also included one model in which *S. iberaensis* was sister to the remaining five species combined. This strategy had the advantage of including a smaller number of parameters than if all species had been included in a single model. Because of the computationally intensive nature of this analysis, each species was represented by 10 individuals, sampling one haplotype per locus per individual. In addition, we only included males in the analysis so that the sampling of the sex chromosome Z would be equivalent to that of autosomes. For the model that combined all species as sister to *S. iberaensis*, we sampled two haplotypes per species (total of 10 haplotypes as in all other models). Each pairwise model estimated 13 demographic parameters: three current and two ancestral population sizes, two splitting times, and six migration rates (*i.e.*, both directions for every pairwise combination among three

populations: two ingroup taxa and one outgroup). In the models that included *S. iberaensis*, we did not allow gene flow between this species and the outgroup *S. bouvreuil*, as these species are allopatric. We ran the multi-threaded version of the program for 750,000 iterations, discarding the initial 150,000 iterations as burn-in, and left the remaining parameters as default. Once each run was completed, we checked that the traces from the different parameter estimates were stationary by visualizing them using the *coda* package in R (Plummer et al. 2006). We converted the median and 95% Bayesian credible intervals for each parameter from mutation scale to generations or individuals using an approximate mutation rate estimate of 10^{-9} per bp per generation (Smeds et al. 2016), as described in Gronau et al. (2011). We expressed migration (*i.e.,* introgression from one species into another) as the number of individual migrants per generation, calculated as indicated in Campagna et al. (2019), which is independent of our assumption of mutation rate. We note that the mutation rate is a rough approximation, and therefore focus the interpretation of the results on relative comparisons between estimates of different parameters, which are not influenced by our assumptions of mutation rate.

Behavioral experiment

To create song stimuli for the behavioral experiment, we recorded the long-range advertisement songs of ten individuals of S. *hypoxantha* and *S. iberaensis*, respectively, in the San Nicolás Reserve during the 2018 breeding season using a Marantz PMD 661 digital recorder at a 48 kHz sampling rate and 16 bit-depth PCM, combined with a Sennheiser ME66 shotgun microphone. For the control files, we downloaded five high-quality tracks of *S. collaris* (catalog numbers: XC47158, XC51947, XC73787, XC108310 and XC108311) that were recorded in the neighboring province of Santa Fe, Argentina from 2006-2012 and archived on Xeno-canto

(https://www.xeno-canto.org/). To generate each playback file, we selected 10-15 of the cleanest songs from each recording, inserted a standard interval of five seconds of silence between them, which is representative of natural singing behavior in capuchinos, and looped over the songs to create a five-minute track. We used Ocenaudio 3.7.9 to standardize the maximum amplitude of all playback files to -3dB and included five seconds of silence at the beginning of each track.

For the plumage stimuli, we created six artificial mounts from hand-painted 3D printed models (two representing *S. hypoxantha*, two representing *S. iberaensis*, and two representing *S. collaris*). We generated a base model of a mounted specimen of a male black-headed grosbeak (*Pheucticus melanocephalus*) at the Denver Museum of Nature and Science using a technique known as photogrammetry, which involves constructing a 3D model from a series of overlapping photographs taken from numerous angles around the specimen. We imported the photographs into the basic version of Agisoft Photoscan, which we used to process the photographs, assemble the model, and generate a high-resolution point cloud. The point cloud was exported in PLY format and loaded into CloudCompare to generate a high-resolution mesh. We then imported the mesh into MeshLab to simplify the model down to one million polygons, used the digital sculpting tool ZBrush to modify the bill and fill in missing parts on the tail and underside of the bird's body, and loaded the model into the sculpting program 3DCoat to make it 'watertight' (*i.e.,* a completely solid object ready for 3D printing). We printed a smaller version of the mount (~9 cm from bill to tail) to represent the two southern capuchino species, as *S. hypoxantha* and *S. iberaensis* are morphometrically indistinguishable, and a larger version of the mount (~12 cm from bill to tail) to represent *S. collaris*, which is considerably larger in size (del Hoyo et al. 2020). The models were printed with a Wanhao Duplicator 5S 3D printer and heavily sanded and painted with an acrylic medium to fill in the uneven surface of the printed models. We used

Vallejo hobby acrylic paints to paint the models and sealed them with a matte varnish (Figure 4.5A-C). Finally, we glued magnets to the base of the models in order to mount them on a pole in the field.

To determine whether the mounts accurately represented male coloration, we referred to standardized photographs and feather samples that were collected from four plumage patches across the body (crown, throat, belly, and rump) of 46 breeding males of *S. hypoxantha* and 41 breeding males of *S. iberaensis* from 2016-2018. Reflectance data for each plumage patch on the live birds were generated following the methods in the *Feather coloration* section, and the colors of the hand-painted models were visually compared to the spectrophotometer readings of *S. hypoxantha* and *S. iberaensis* breeding males to ensure accurate coloration (Figure S4.14). Both capuchino species have crown feathers that reflect wavelengths in the near ultraviolet (300-400 nm), which could not be incorporated into the models. Nonetheless, live birds of *S. hypoxantha* and *S. iberaensis* do not differ in crown coloration. As a result, failing to incorporate UV reflectance into the models should have the same effect on the behavioral responses of both species. Each of the six mounts (including the two heterospecific controls) elicited attack behavior from conspecific males over the course of the study, which suggests that the mounts were sufficiently realistic representations of the three species.

SUPPLEMENTAL TEXT

Natural history of S. iberaensis and history of the San Nicolás Reserve and museum collections in Iberá National Park

Until recently, the fauna of Iberá National Park was primarily represented in museum collections from localities in the south (*e.g.,* near Colonia Carlos Pellegrini; 28° 32' 10" S, 57° 10' 25" W).

Ornithological expeditions, including those by William H. Partridge in the 1960s, and Julio R. Contreras in the 1990s (specimens deposited in the Museo Argentino de Ciencias Naturales "Bernardino Rivadavia" in Argentina and museums in the United States), generally focused on other areas of the province of Corrientes. Indeed, many capuchinos that are common in Iberá National Park first appeared in collections in the 1970s (Short 1971), and common grassland birds from Corrientes are poorly represented relative to birds from other regions (Collar 1992). Giraudo et al. (2003) and Fraga (2001) explored several areas in the northern portion of the Esteros del Iberá wetlands (Estancia El Tránsito, 28° 29' S, 57° 40' W; Estancia San Nicolás, 28° 8' S, 57° 26' W; Loreto and Estancia San Juan Poriahú; 27° 42' S, 57° 11' W), but did not report *S. iberaensis* or *S. hypoxantha*, now the two most abundant capuchino species in San Nicolás. The San Nicolás Reserve served as a cattle farm until 2009, when it was converted into a protected area, possibly contributing to grassland recovery and the population increase of *S. iberaensis*.

Haplotype phasing from the divergence peaks on scaffolds 257 and 430

We phased the variants found within the peak on scaffold 430 to obtain the haplotypes observed among the 37 individuals (16 *S. hypoxantha* and 21 *S. iberaensis*). This region contained 157 SNPs with an F_{ST} value above 0.5 (mean of 0.3209 ± 0.2817 SD for a total of 625 SNPs), allowing most haplotypes from each species to cluster together (Figure S4.6; with the exception of one *S. hypoxantha* haplotype from a male with typical plumage). However, when we conducted the same analysis on the 64 SNPs showing the highest level of differentiation within the peak, we observed two haplotypes of *S. iberaensis* that clustered with those of *S. hypoxantha* (Figure S4.7). The equivalent plot from the peak on scaffold 257 (derived from 13 SNPs) showed six *S. iberaensis* haplotypes that clustered with those of *S. hypoxantha* (Figure S4.8).

One *S. iberaensis* male had a mismatched haplotype on both peaks, yet possessed *S. iberaensis* breeding plumage.

The different haplotype-based analyses from these peak regions are consistent with incomplete lineage sorting or low levels of hybridization. This result is also supported by the extensive sharing of mitochondrial DNA between *S. iberaensis* and *S. hypoxantha* (Figure S4.3). We note, however, that we did not observe hybridization directly in the pairs sampled in this study, nor individuals that could clearly be assigned as early generation hybrids. In addition, these findings indicate potential redundancy in the genes contributing to phenotypic differences across peaks and suggest that the genotypes that lead to *S. iberaensis* plumage may be dominant over those that produce *S. hypoxantha* plumage, as *S. iberaensis* males with mixed genotypes in the variants showing the highest level of differentiation still possess *S. iberaensis* plumage. A better understanding of the causal sites responsible for both plumage patterns is needed to shed light on how *TYRP1* and *OCA2* act together to shape these phenotypes.

Demographic modeling and relationships between S. iberaensis and other southern capuchinos

Even though *S. iberaensis* is broadly sympatric with other southern capuchinos in Iberá National Park, and breeds side by side with *S. hypoxantha* in the San Nicolás Reserve, *S. iberaensis* individuals form their own clade in a phylogenetic tree (Figure 4.6A) and can be clearly differentiated from other capuchino seedeaters in PCAs based on SNPs from either whole-genome or ddRAD sequencing data (Figure S4.18A-B). While the *S. iberaensis* clade is sister to a clade composed mostly of *S. hypoxantha* individuals (Figure 4.6A), the relationship between *S. iberaensis* and other species in the radiation remains unclear. To better understand the evolutionary history of *S. iberaensis*, we conducted demographic analyses in G-PhoCS, which allowed us to infer splitting times, effective population sizes, and migration rates between *S.*

iberaensis and five other southern capuchinos (*S. hypoxantha*, *S. melanogaster*, *S. pileata*, *S. ruficollis* and *S. palustris*). We also compared the parameters estimated for *S. iberaensis* mentioned above to equivalent parameters inferred between the latter southern capuchino species alone (*i.e.,* without the inclusion of *S. iberaensis*). We conducted a total of 16 G-PhoCS runs, each including one of the 15 pairwise combinations of six capuchino species as sister to each other and *S. bouvreuil* as an outgroup (Campagna et al. 2013). The 16th G-PhoCS run compared *S. iberaensis* to the five remaining southern capuchinos combined, with *S. bouvreuil* as an outgroup (Figure S4.18C).

The results from the 16 G-PhoCS runs could be grouped into two general patterns. The majority of runs (11 out of 16) identified a divergence time within the ingroup comparable to that between the ingroup and *S. bouvreuil*, and inferred a smaller ancestral effective population size than the current effective population sizes of the ingroup taxa (Figure S4.18D). These models identified migration between the three current populations, including from and into *S. iberaensis*. The second pattern (Figure S4.18E) included 5 models (S. *hypoxantha* vs. *S. iberaensis*, *S. ruficollis* vs. *S. iberaensis*, *S. hypoxantha* vs. *S. melanogaster*, *S. palustris* vs. *S. pileata*, and *S. palustris* vs. *S. ruficollis*) that identified a very recent split between the ingroup taxa and an older split between the ingroup and the outgroup. These models estimated a large ancestral effective population size and smaller current effective population sizes for the ingroup, and detected gene flow only between the ingroup and the outgroup. In addition, these models include the two comparisons that show the lowest mean genome-wide F_{ST} with respect to *S. iberaensis*: *S. hypoxantha* (0.0055 ± 0.0137 SD), and *S. ruficollis* (0.0037 ± 0.0184 SD; Figure S4.20). We note that these estimates are based on a rough assumption of mutation rate and do not include all species interacting in the model at the same time. Therefore, the results must be

interpreted with caution. However, they contribute to our understanding of the evolutionary history of *S. iberaensis* by showing that, although this species was recently discovered, it has similar demographic parameters to those observed among the other southern capuchinos. For example, the most recent identified split between two capuchinos was estimated at ~1500 generations, between *S. iberaensis* and *S. hypoxantha*. However, the 95% credible interval of this estimate overlapped with the inferred splitting time between *S. hypoxantha* and *S. melanogaster*. We also observed levels of gene flow between *S. iberaensis* and other species (*S. melanogaster*, *S. pileata*, and *S. palustris*) that are similar to those inferred between other capuchinos, indicating low levels of hybridization during speciation. These results suggest that *S. iberaensis* could have diverged recently from *S. hypoxantha*, and it is possible that the process occurred with the introgression of specific genomic regions from other taxa.

SUPPLEMENTAL FIGURES

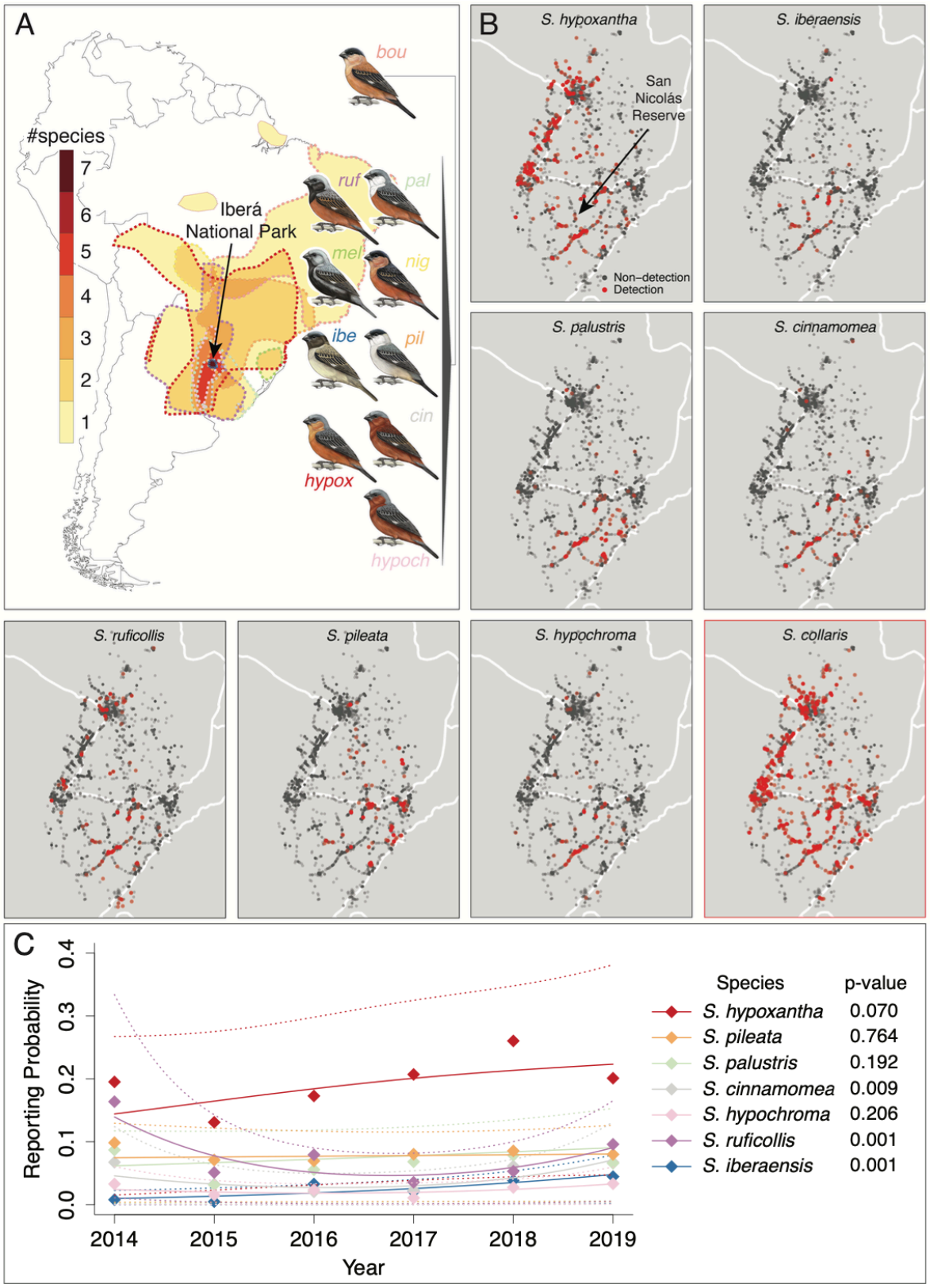


Figure S4.1. Geographic distribution of the southern capuchino seedeaters and their prevalence in Iberá National Park. **(A)** Cartoon showing the range overlap of southern capuchino seedeaters, modified from (Campagna et al. 2017). Species are abbreviated using the first few letters of each name. **(B)** Location of eBird reports across the breeding range of *S. iberaensis* for the seven capuchino species (rectangles outlined in gray) that breed in Iberá National Park, as well as *S. collaris* (lower right corner; rectangle outlined in red), the control in the behavioral experiment. The red circles indicate locations where each focal species was observed, and the gray symbols denote areas where eBird checklists were submitted but the focal species was not detected for years in which the focal species was reported in the eBird database. Across the limited breeding range of *S. iberaensis*, *S. hypoxantha* is the most common and *S. iberaensis* is one of the least common capuchino species. *S. collaris*, while not a member of the southern capuchino radiation, is common in Iberá National Park and breeds in sympatry with *S. iberaensis* and *S. hypoxantha*. **(C)** Change in reporting probability in the eBird database from 2014-2019 for the capuchino species that breed in Iberá National Park. The points depict reporting probability per year and the solid lines indicate the fitted results of generalized additive models (GAMs) that model the effect of year on reporting probability while controlling for the confounding effects of observation date, distance traveled, and duration of the observation period. The dotted lines show the 95% prediction intervals for the estimated probabilities. The p-values from the GAMs are the probabilities that there have been systematic changes in the relationship between year and reporting probability for each capuchino species. *S. iberaensis* has increased in local prevalence to surpass that of *S. hypochroma* across its limited breeding range. *S. ruficollis* and *S. cinnamomea* also showed systematic changes in reporting probability, but did not consistently increase in prevalence over time.

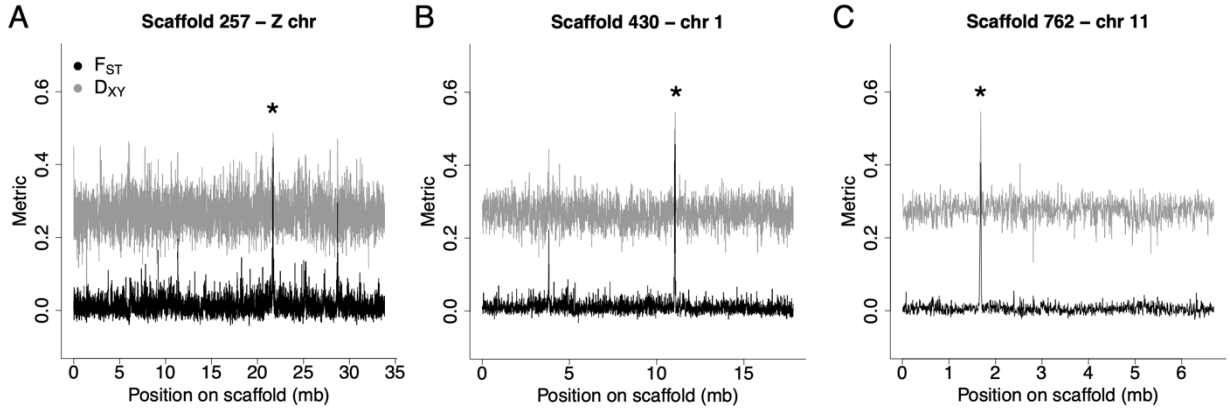


Figure S4.2. F_{ST} peaks were characterized by increased absolute differentiation between *S. hypoxantha* and *S. iberaensis*. Plots comparing patterns of relative genomic differentiation (F_{ST}) and absolute genomic differentiation (D_{XY}) over 5-kb windows across **(A)** scaffold 257, **(B)** scaffold 430, and **(C)** scaffold 762 ($N = 37$). Asterisks indicate the divergence peaks that contained at least one individual SNP with F_{ST} greater than 0.85 (as in Figure 4.1F).

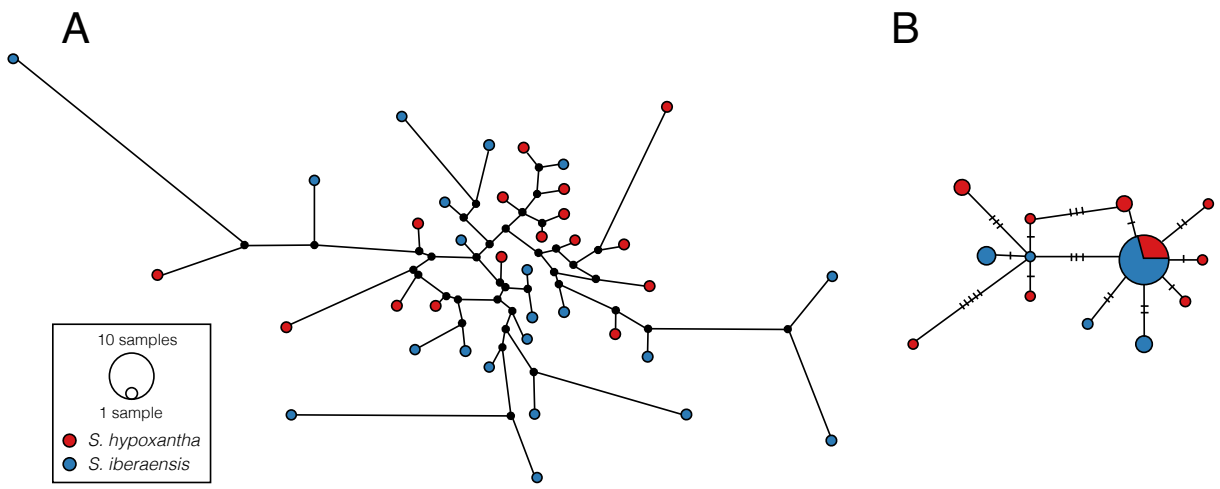


Figure S4.3. Mitochondrial unrooted statistical parsimony networks for *S. hypoxantha* and *S. iberaensis*. **(A)** Haplotype network based on a 17,004 bp alignment of the mitochondrial genome. **(B)** Haplotype network based on 697 bp of the mitochondrial COI DNA barcode region commonly used for species identification. The length of the branches connecting haplotypes is proportional to the number of nucleotide differences between them, which are indicated by the number of line marks in each branch (omitted for simplicity in the case of the full mitochondrial network). Black dots in (A) indicate hypothetical, unobserved haplotypes. The two capuchino species show no differentiation in mitochondrial DNA with either approach.

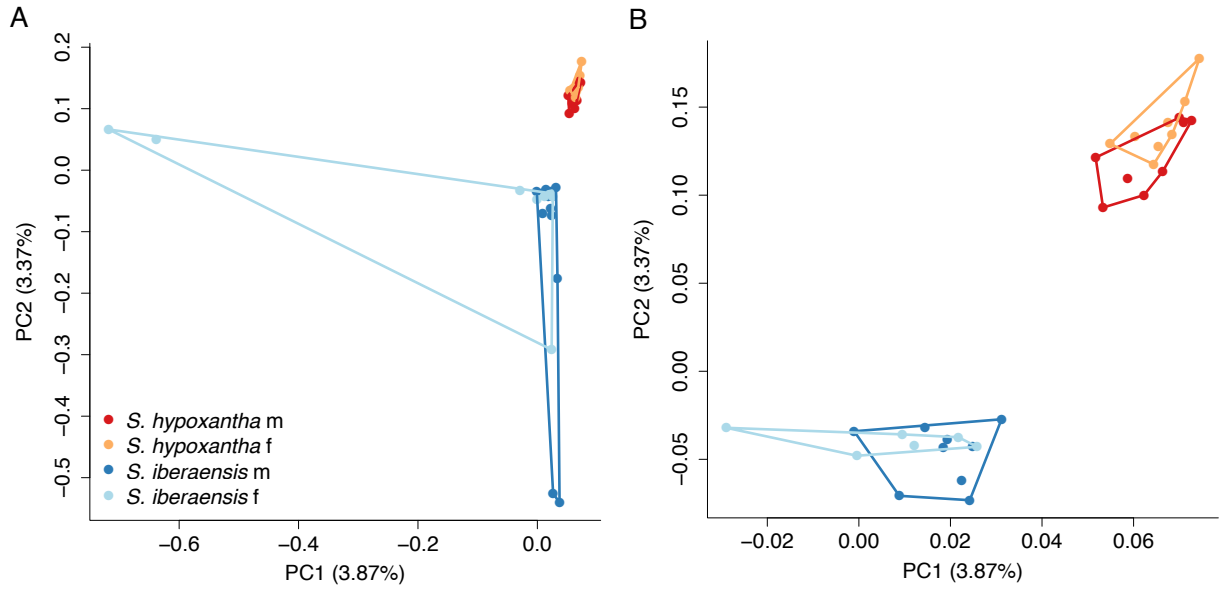


Figure S4.4. Pattern of genomic differentiation between *S. hypoxantha* and *S. iberaensis*. Principal component analysis (PCA) generated with whole-genome data (~13.3 million SNPs) depicting the degree of genomic differentiation between **(A)** all males and females of *S. hypoxantha* and *S. iberaensis* ($N = 37$) and **(B)** individuals with a low relatedness statistic ($N = 31$). Three pairs of *S. iberaensis* individuals, which appear as outliers in (A), were more closely related to each other than the others (unadjusted A_{jk} relatedness statistic: B009595 vs. B009507 =

0.46; B009502 vs. B009504 = 0.19; B009508 vs. B009542 = 0.07). Female capuchinos were classified based on the phenotype of their social mate.

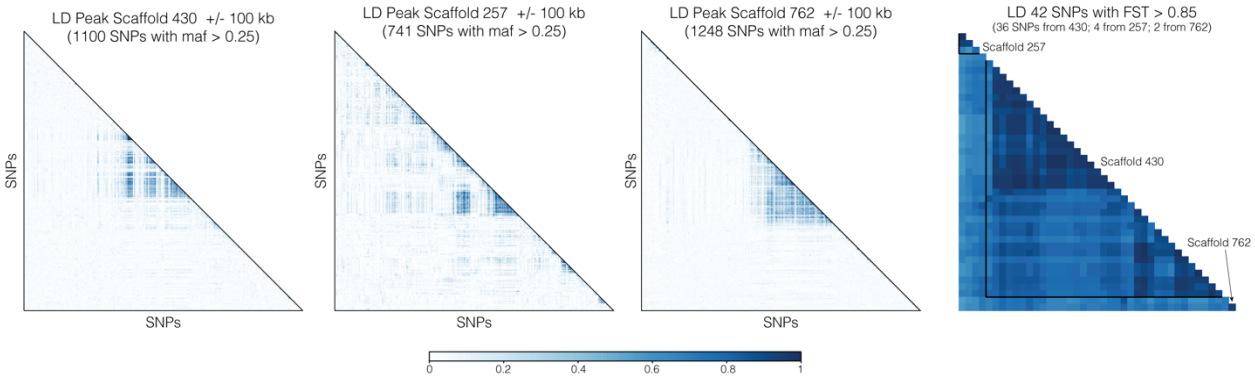


Figure S4.5. Pattern of linkage disequilibrium (LD). LD measured by the r^2 statistic across the peak on scaffold 430, 257, and 762, as well as among all possible combinations of the 42 highly differentiated SNPs with $F_{ST} > 0.85$ within and among each of the three peaks. LD between SNPs located within each peak was high (mean: 0.82, range: 0.59-1 for the peak on scaffold 257; mean: 0.82, range 0.56-1 for the peak on scaffold 430; and 0.94 for the comparison of the two SNPs on the peak on scaffold 762). The values of LD among comparisons of SNPs from different peaks (which were located on separate chromosomes) were also elevated, with a mean of 0.69 (0.56-0.88). Inter-chromosomal LD was somewhat lower but comparable when calculated separately for *S. hypoxantha* (mean: 0.26) and *S. iberaensis* (mean: 0.28). However, the two species tend to have different alleles at each of the 42 highly differentiated SNPs within the three divergence peaks (see haplotype plots below).


SNPs from the peak on scaffold 430
S. hypoxantha
S. iberaensis
C
T
G
A

Figure S4.6. Nearly all individuals cluster by species in a haplotype tree derived from the SNPs within the peak on scaffold 430. Each row represents a single chromosome, and each individual is represented twice in the tree. The analysis was conducted with the individuals for which we had whole-genome data (*S. hypoxantha*: $N = 16$; *S. iberaensis*: $N = 21$). The four nucleotides are color-coded as indicated in the upper left corner. One haplotype from *S. hypoxantha* male B009525 grouped with those of *S. iberaensis*.

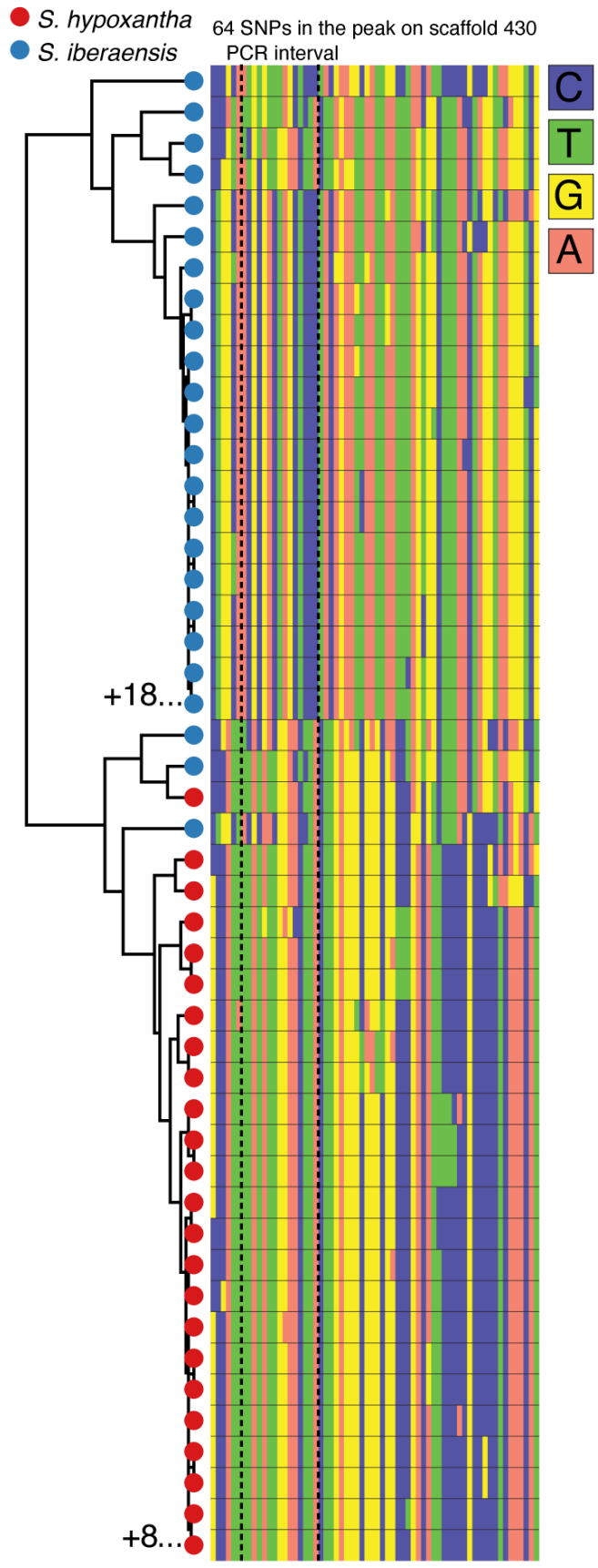


Figure S4.7. Clustering of haplotypes from the SNPs showing the highest F_{ST} in the peak on scaffold 430. Most haplotypes (two per individual: *S. hypoxantha*, $N = 32$; S. *iberaensis*, $N = 42$) cluster by species, except for three haplotypes derived from *S. iberaensis* that group with those of *S. hypoxantha* (one haplotype from male B009552, one from male B009504, and one from female B009507). The analysis included the 36 SNPs with $F_{ST} > 0.85$ in this region, plus an additional 28 SNPs with $F_{ST} > 0.79$. We included these additional SNPs as the segment selected for PCR amplification (indicated by vertical dashed lines) included three SNPs with this level of divergence. For graphical clarity, 18 identical copies of the most common *S. iberaensis* haplotype and 8 identical copies of the most common *S. hypoxantha* haplotype were omitted from the tree. Other details as in Figure S4.6.

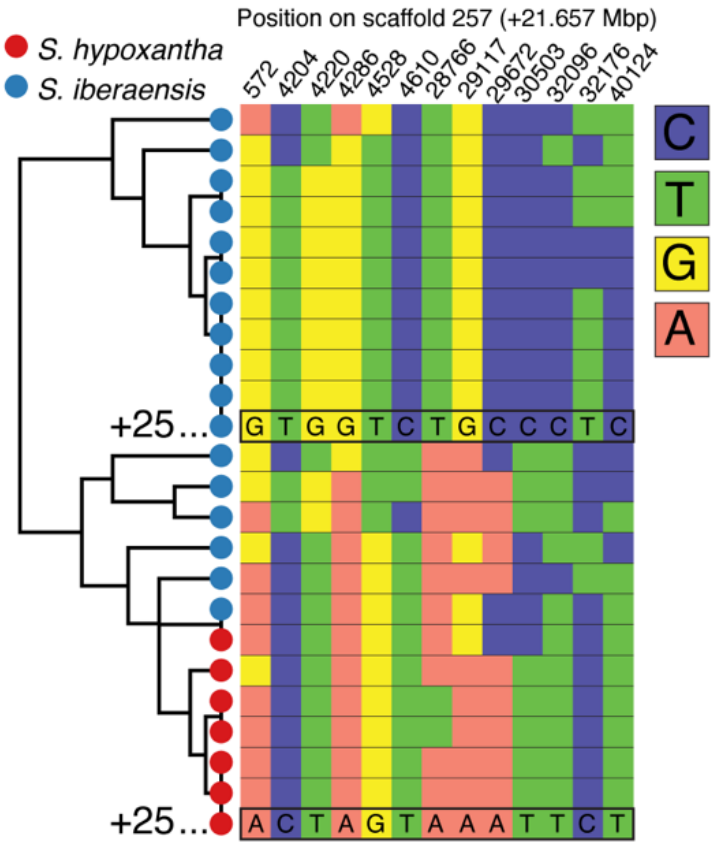


Figure S4.8. Clustering of haplotypes from the SNPs showing the highest F_{ST} in the peak on scaffold 257. The haplotypes from *S. hypoxantha* ($N = 32$) and *S. iberaensis* ($N = 42$) individuals cluster by species, except for six *S. iberaensis* haplotypes that group with those of *S. hypoxantha*. The mismatched haplotypes belong to male B009508 (both haplotypes), male B009561 (both haplotypes), B009552 (one haplotype), and B009583 (one haplotype). Note that male B009552 also possesses a mismatched haplotype in the peak on scaffold 430 (Figure S4.7), yet has typical *S. iberaensis* plumage. This analysis included the 13 SNPs with $F_{ST} > 0.79$ in this region. The most common haplotype for each species is indicated at the bottom of the two main clades, and 25 identical copies of these haplotypes from each species were omitted for graphical clarity. Other details as in Figure S4.6.

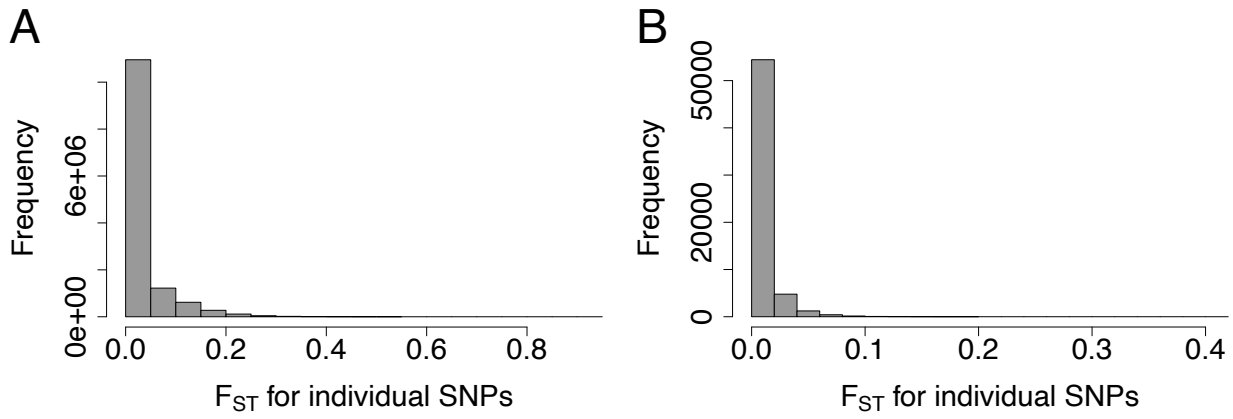


Figure S4.9. Few SNPs with F_{ST} greater than 0.2 were present in the whole-genome and ddRAD datasets. Histograms showing the frequency of F_{ST} values for individual SNPs in the **(A)** whole-genome and **(B)** ddRAD data. Only 1.4% of SNPs in the whole-genome data and 0.008% of SNPs in the ddRAD data had moderate F_{ST} values (> 0.2).

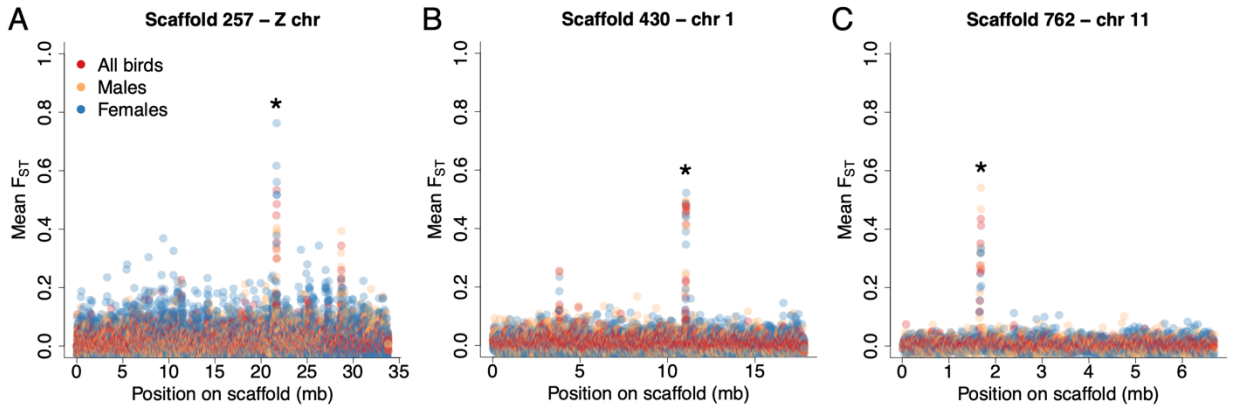


Figure S4.10. Females and males exhibit elevated divergence in the same genomic regions. Manhattan plots showing regions of elevated genomic differentiation (F_{ST}) over 5-kb windows on **(A)** scaffold 257, **(B)** scaffold 430, and **(C)** scaffold 762 for all individuals ($N = 37$), males ($N = 20$), and females ($N = 17$) of *S. hypoxantha* and *S. iberaensis*. Asterisks indicate the divergence peaks that contained at least one individual SNP with F_{ST} greater than 0.85 when all individuals were compared (as in Figure 4.1F). Female capuchinos were classified based on the phenotype of their social mate.

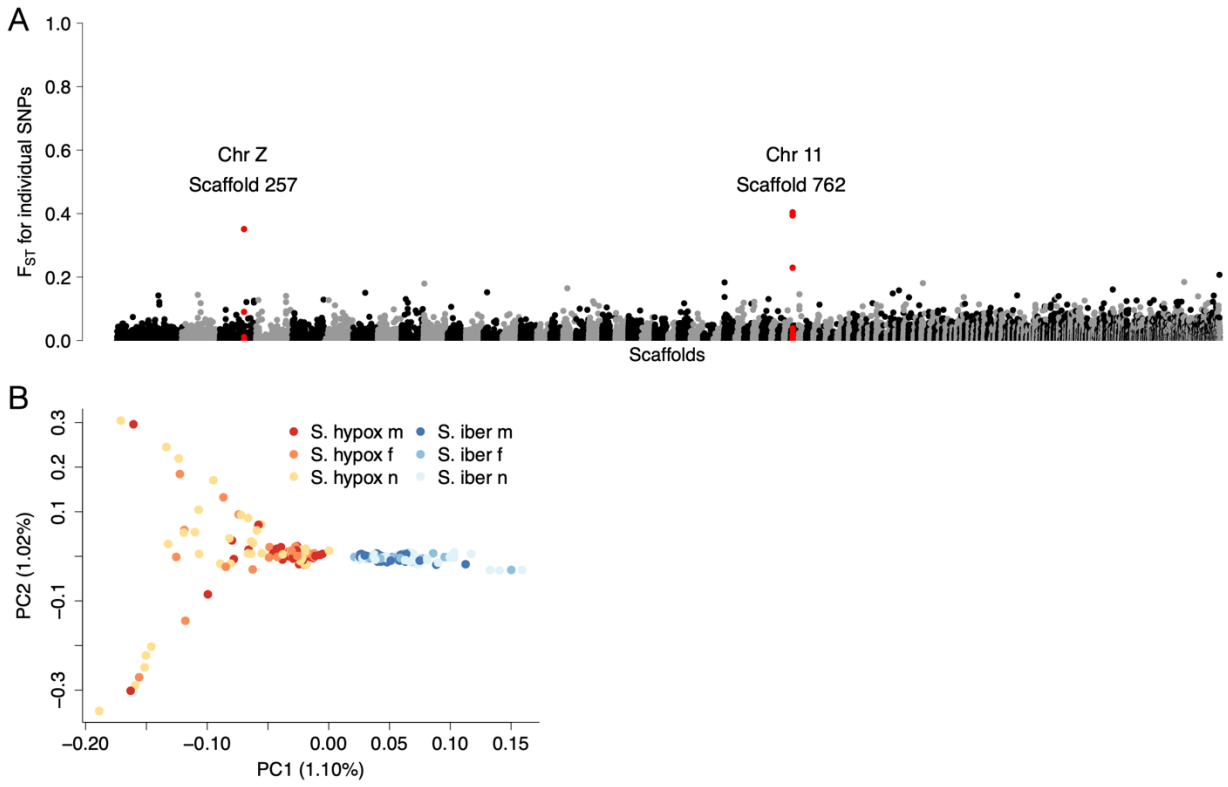


Figure S4.11. Few SNPs in the ddRAD data fell within the F_{ST} peaks found in the whole-genome data. **(A)** Manhattan plot generated with ddRAD data showing the pattern of genomic differentiation between *S. hypoxantha* and *S. iberaensis* ($N = 206$). The 28 SNPs that fell within F_{ST} peaks identified from the whole-genome data are highlighted in red and labeled according to their scaffold and corresponding chromosome in the zebra finch assembly. The plot contains 571 scaffolds. **(B)** PCA generated with ddRAD data depicting the pattern of genomic differentiation between males, females, and nestlings of *S. hypoxantha* and *S. iberaensis* when the 28 SNPs that fell within F_{ST} peaks identified from the whole-genome data were removed ($N = 206$ individuals). The ddRAD data distinguished the two species even when the SNPs that fell within F_{ST} peaks were excluded. The two species are separated in PC1, while *S. hypoxantha* shows additional variation in PC2.

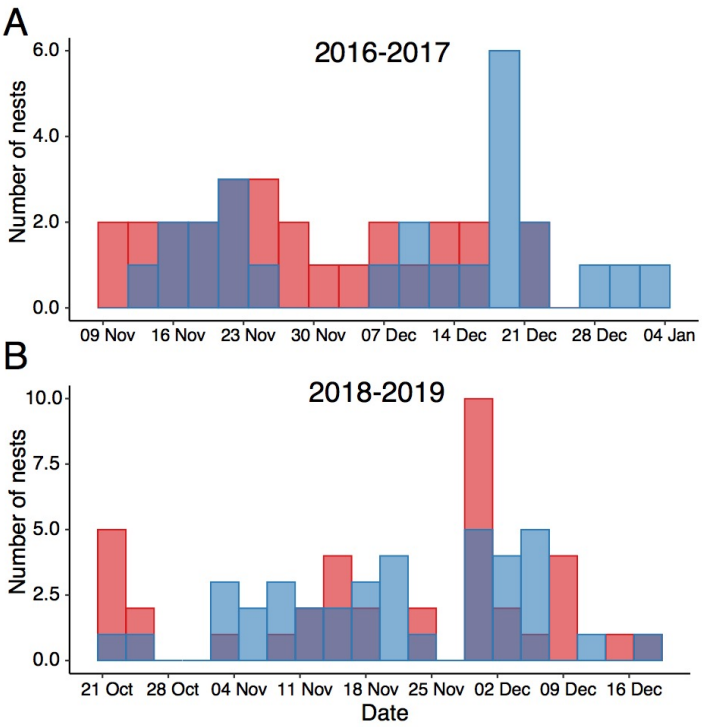


Figure S4.12. *S. hypoxantha* and *S. iberaensis* breed synchronously in Iberá National Park. Histograms showing the number of nests of *S. hypoxantha* (red) and *S. iberaensis* (blue) found in the San Nicolás Reserve over the course of two breeding seasons: **(A)** 2016-2017 ($N = 52$) and **(B)** 2018-2019 ($N = 76$). Overlapping bars are shown in purple.

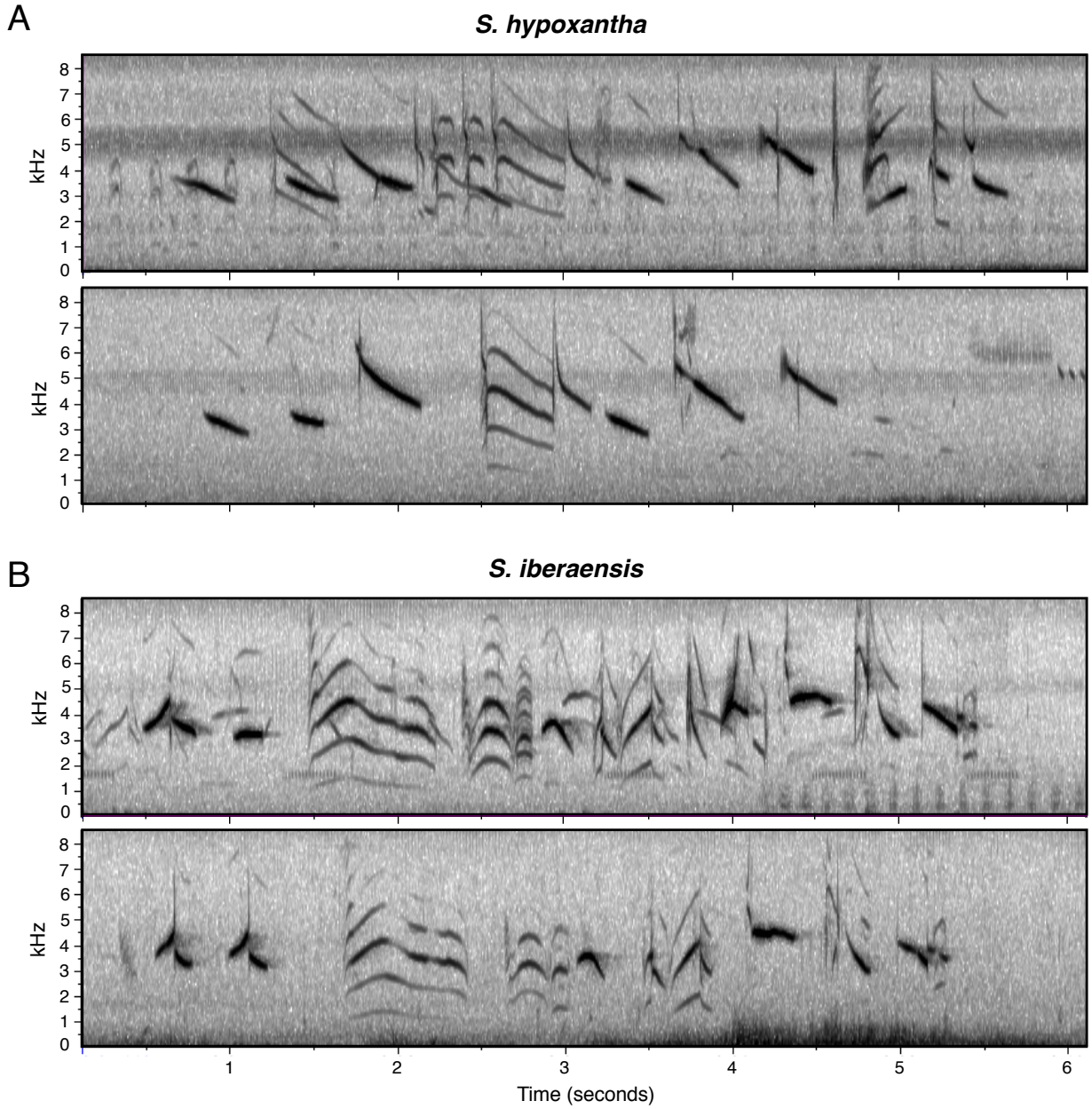


Figure S4.13. Differences in male song between *S. hypoxantha* and *S. iberaensis.* Representative spectrograms showing the typical song of two **(A)** *S. hypoxantha* and **(B)** *S. iberaensis* males.

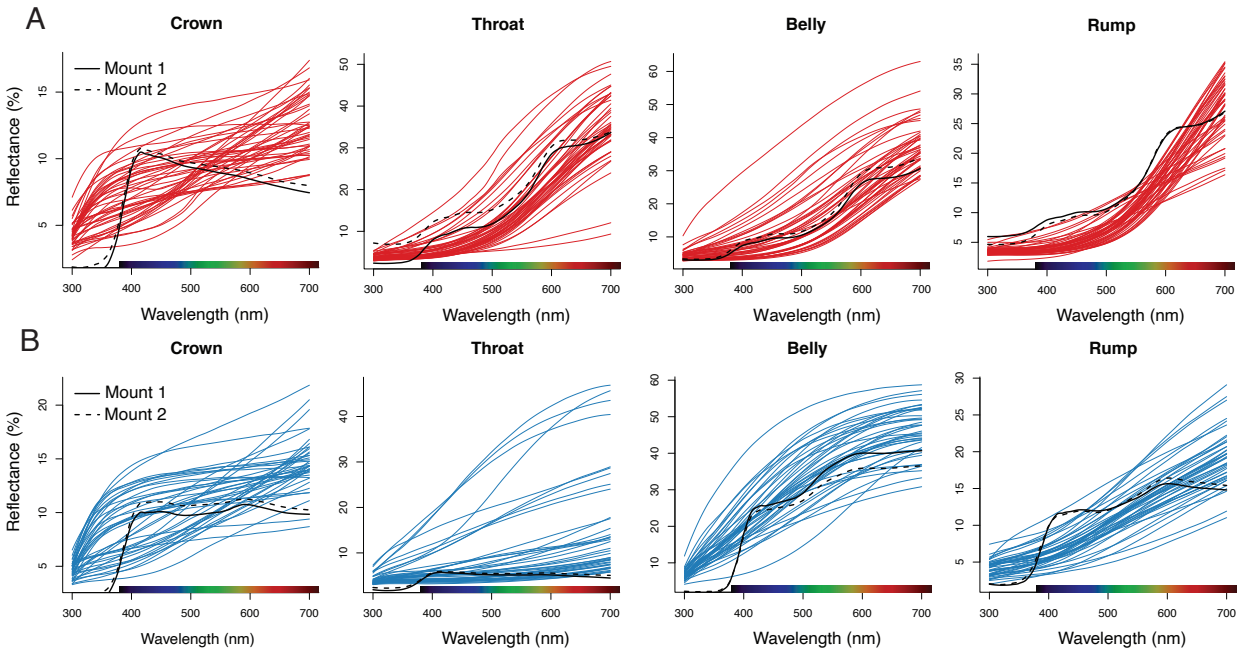
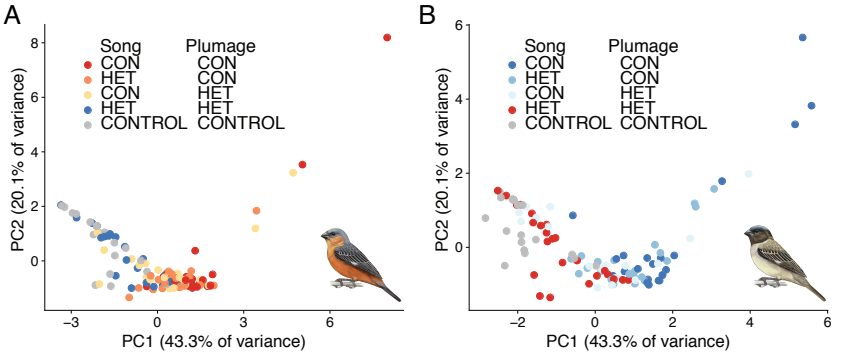


Figure S4.14. Comparison between the plumage coloration of breeding males and the mounts used in the behavioral experiment. Reflectance patterns across the avian visual spectrum for the crown, throat, belly, and rump of the two mounts used for each species (solid black and dashed lines) in the behavioral experiment and breeding males of **(A)** *S. hypoxantha* ($N = 46$) and **(B)** *S. iberaensis* ($N = 41$).



C

Behavioral Response	PC1	PC2
Percent variation	43%	20%
Flights under 6 m	0.50	-0.01
Attacks at the mount	0.39	0.50
Proportion time under 3 m	0.48	-0.32
Proportion time 3-6 m	-0.12	-0.27
Proportion time over 6 m	-0.43	0.47
Proportion time attacking	0.31	0.57
Proportion time singing	-0.26	0.16

Figure S4.15. Principal component analysis (PCA) of the correlation matrix of the behavioral response variables. PC axes for territorial males of **(A)** *S. hypoxantha* ($N = 120$) and **(B)** *S. iberaensis* ($N = 120$) when presented with various combinations of conspecific (CON), heterospecific (HET), and *S. collaris* (CONTROL) song and plumage. Points are colored by treatment. **(C)** Behavioral responses and their factor loadings for the first two principal components. Positive values of PC1 were associated with more flights within six meters of the mount, more attacks at the mount, and a greater proportion of time spent within three meters of mount.

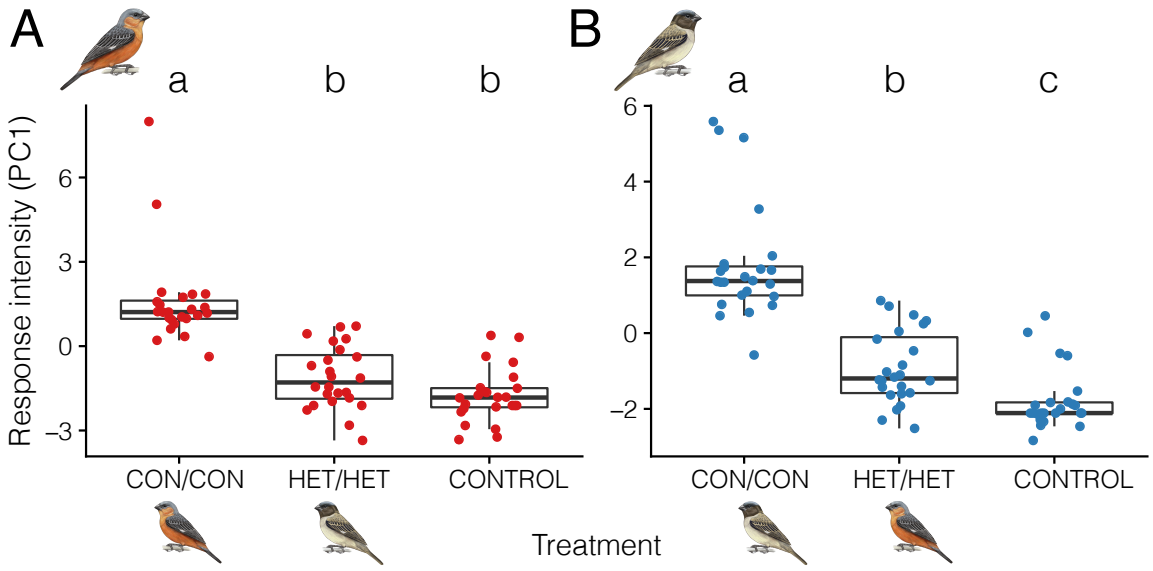


Figure S4.16. Territorial males of **(A)** *S. hypoxantha* and **(B)** *S. iberaensis* responded more aggressively to conspecific song and plumage (CON/CON) than that of the heterospecific capuchino (HET/HET) or control (*S. collaris*; CONTROL). Points are colored by species and positive values of PC1 were associated with more flights within six meters of the mount, more attacks at the mount, and a greater proportion of time spent within three meters of mount ($N = 24$ trials per treatment). We ran generalized linear mixed models examining the effect of treatment group on response intensity when the mismatched treatments were excluded. Different letters indicate statistical significance between treatment groups (Tukey HSD method; adjusted $P < 0.05$). For *S. hypoxantha*, we detected significant differences between the CON/CON and HET/HET treatments ($P < 0.0001$) and the CON/CON and CONTROL treatments ($P < 0.0001$). For *S. iberaensis*, we detected significant differences between all treatment groups (CON/CON vs. HET/HET: $P < 0.0001$, CON/CON vs. CONTROL: $P < 0.0001$, HET/HET vs. CONTROL: $P = 0.008$). When outliers (observations outside 1.5 * interquartile range; 7 for *S. hypoxantha* and 10 for *S. iberaensis*) were removed, we detected an additional significant difference in the response of *S. hypoxantha* to the HET/HET and CONTROL treatments ($P = 0.003$).

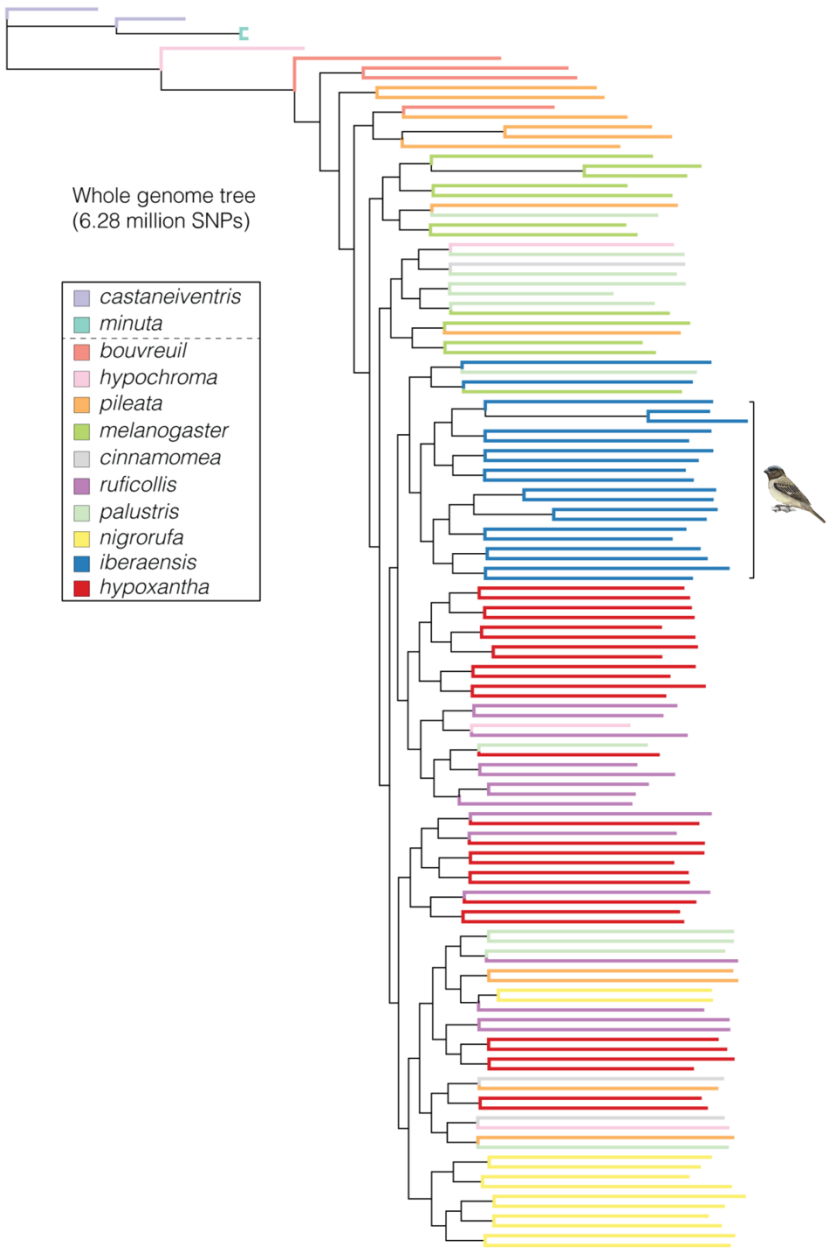


Figure S4.17. Phylogenetic tree inferred using maximum likelihood depicting the relationship between ten capuchino species and two outgroups based on whole-genome data. This dataset retained 6,283,771 SNPs that were present in at least 85% of individuals and had a minor allele frequency of at least 10%. The black square bracket to the right of the tips in the tree indicates the clade containing the majority of *S. iberaensis* individuals. Outgroups are indicated above the dashed line in the legend.

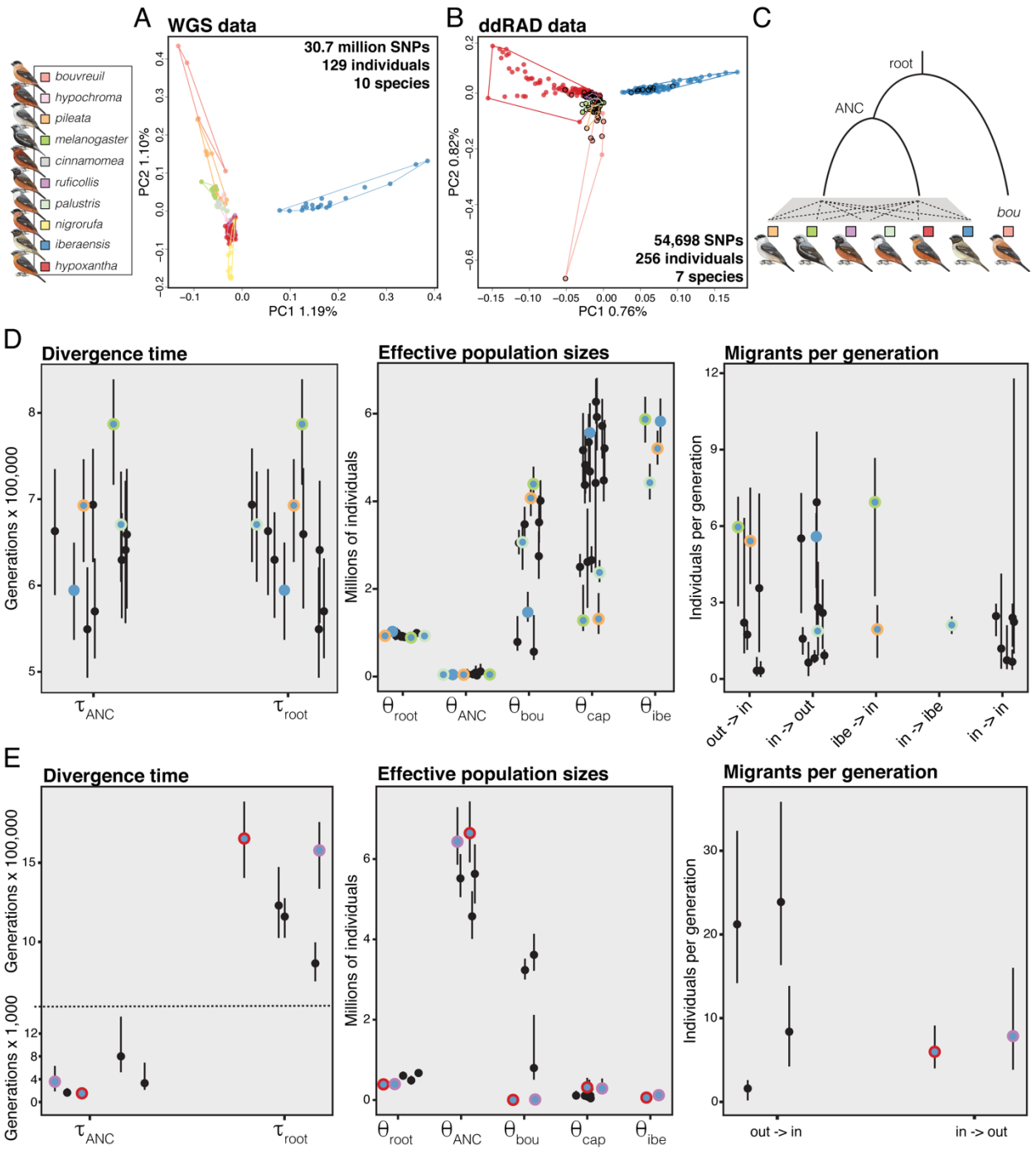


Figure S4.18. Demographic reconstructions for southern capuchino seedeaters. Clustering patterns in PCAs for capuchino individuals based on **(A)** whole-genome sequencing (WGS) data with the individuals and species shown in Figure 4.6, and **(B)** ddRAD data. For the PCA based on WGS data, we excluded the outgroups *S. minuta* and *S. castaneiventris*, one outlier *S. hypochroma* individual, and three other individuals showing high levels of relatedness with respect to another member of their species (one *S. iberaensis*, one *S. pileata*, and one *S. melanogaster*). The individuals used for demographic modeling are shown in (B) as circles with a black outline. **(C)** Tree representing the ancestral and current populations used for demographic

reconstructions. The results from the 16 G-PhoCS runs could be grouped into two general patterns. (**D**) The first pattern, which included most runs, identified similar divergence times within the ingroup and between the ingroup and *S. bouvreuil*, inferred a smaller ancestral effective population size than the current effective population sizes of the ingroup taxa, and detected gene flow among the three populations. **(E)** The second pattern included five runs that identified a much more recent split between the ingroup taxa, estimated smaller current effective population sizes for the ingroup, and only detected gene flow with the outgroup. For each run, the posterior median value is shown with a circle and the 95% credible interval with a vertical black line. Blue circles represent estimates from models that included *S. iberaensis*, with the second ingroup species indicated by the color of the outline. The solid blue circles represent the comparison of *S. iberaensis* vs. all other ingroup species combined. *S. bouvreuil* is abbreviated as bou, capuchinos as cap, and *S. iberaensis* as ibe. Note the split y-axis for the divergence time plot in (E). Only estimates of migration with 95% credible intervals that do not overlap with zero are shown. The two highest migration estimates from the ingroup to the outgroup in (E) are not shown in the plot due to their large 95% credible interval (ranging from 61 to 343 and 6 to 97, respectively). The split between *S. iberaensis* and *S. hypoxantha*, estimated at 1500 generations, was the most recent split identified between two capuchinos.

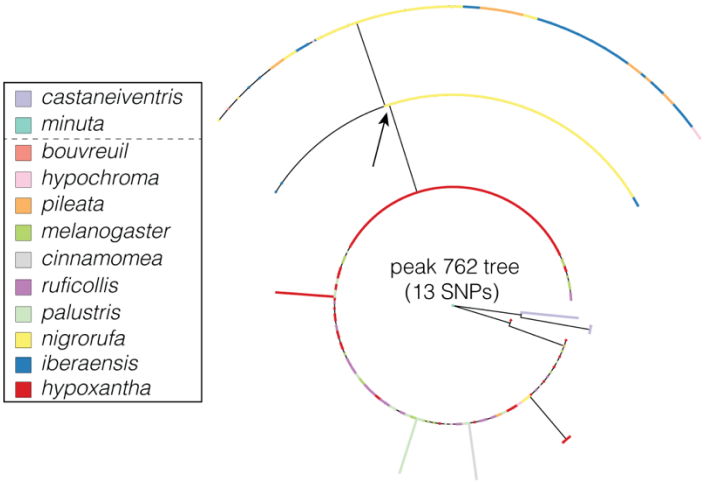


Figure S4.19. Phylogenetic tree inferred using maximum likelihood depicting the relationship between ten capuchino species and two outgroups based on the 13 SNPs that fell within the F_{ST} peak on scaffold 762. The tree was generated using whole-genome data and the arrow indicates the clade that contains all individuals of *S. iberaensis*, as well as many other species.

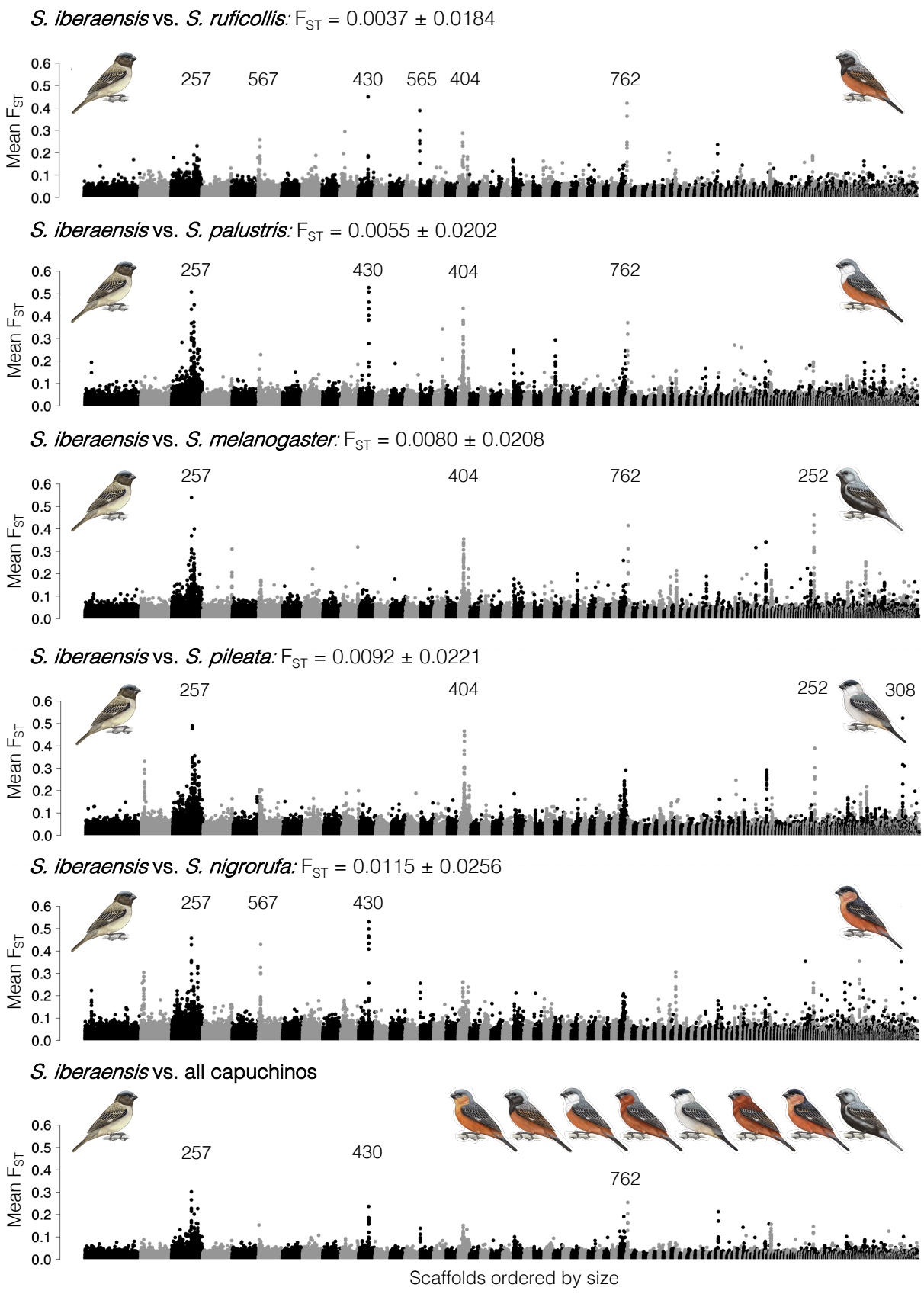

S. iberaensis vs. *S. ruficollis*: F_{ST} = 0.0037 ± 0.0184
257
567
430
565
404
762
S. iberaensis vs. *S. palustris*: F_{ST} = 0.0055 ± 0.0202
257
430
404
762
S. iberaensis vs. *S. melanogaster*: F_{ST} = 0.0080 ± 0.0208
257
404
762
252
S. iberaensis vs. *S. pileata*: F_{ST} = 0.0092 ± 0.0221
257
404
252
308
S. iberaensis vs. *S. nigrorufa*: F_{ST} = 0.0115 ± 0.0256
257
567
430
S. iberaensis vs. all capuchinos
257
430
762
Mean F_{ST}
0.0
0.1
0.2
0.3
0.4
0.5
0.6
Scaffolds ordered by size

Figure S4.20. Manhattan plots from the whole-genome data showing regions of elevated genomic differentiation (F_{ST}) over 5-kb windows between *S. iberaensis* and other southern capuchino species. Genome-wide F_{ST} (mean ± SD) is indicated for each comparison and the main peaks are labeled with their corresponding scaffold, which matched the following chromosomes in the zebra finch assembly: Chr 1 (scaffold 430), Chr 2 (scaffold 567), Chr 4 (scaffold 565), Chr 11 (scaffold 762), Chr 20 (scaffold 252), and Chr Z (scaffolds 257, 404, and 308). The plots contain the 381 largest scaffolds. The final comparison examined differentiation between *S. iberaensis* and eight other capuchino species combined (*N* = 21 *S. iberaensis*, 28 *S. hypoxantha*, 15 *S. ruficollis*, 12 *S. pileata*, 12 *S. palustris*, 12 *S. melanogaster*, 12 *S. nigrorufa*, 3 *S. hypochroma*, and 3 *S. cinnamomea*). The additional divergence peaks contain *SLC45A2* (scaffold 404) and *ASIP* (scaffold 252), two genes involved in the melanogenesis pathway (Poelstra et al. 2015).

SUPPLEMENTAL TABLES

Table S4.1. Individuals sampled for the study and the type of genetic data obtained for each individual. We generated whole-genome sequencing (WGS) data for 37 individuals of *S. hypoxantha* (*N* = 8 males, 8 females) and *S. iberaensis* (*N* = 12 males, 9 females) and double-digest restriction site-associated DNA (ddRAD) sequencing data for 206 individuals of *S. hypoxantha* (*N* = 40 males, 23 females, 40 nestlings, and 2 unhatched eggs) and *S. iberaensis* (*N* = 42 males, 21 females, 37 nestlings, and 1 unhatched egg). The Sanger column indicates the individuals for which we obtained Sanger sequences for a region in the peak on scaffold 430 (Figure 4.2).

Species	Sex	Band No.	Year	WGS	ddRAD	Sanger
S. hypoxantha	Male	B009522	2016	Y	Y	N
S. hypoxantha	Male	B009524	2016	Y	Y	N
S. hypoxantha	Male	B009525	2016	Y	Y	N
S. hypoxantha	Male	B009526	2016	Y	Y	N
S. hypoxantha	Male	B009530	2016	Y	Y	N
S. hypoxantha	Male	B009554	2016	Y	Y	N
S. hypoxantha	Male	B009556	2016	Y	Y	N
S. hypoxantha	Male	B009574	2016	Y	Y	N
S. hypoxantha	Female	B009503	2016	Y	Y	N
S. hypoxantha	Female	B009523	2016	Y	Y	N
S. hypoxantha	Female	B009529	2016	Y	Y	N
S. hypoxantha	Female	B009555	2016	Y	Y	N
S. hypoxantha	Female	B009557	2016	Y	Y	N
S. hypoxantha	Female	B009570	2016	Y	Y	N
S. hypoxantha	Female	B009573	2016	Y	Y	N
S. hypoxantha	Female	B009580	2016	Y	Y	N
S. iberaensis	Male	B009502	2016	Y	Y	N
S. iberaensis	Male	B009504	2016	Y	Y	N
S. iberaensis	Male	B009508	2016	Y	Y	N
S. iberaensis	Male	B009528	2016	Y	Y	N
S. iberaensis	Male	B009540	2016	Y	Y	N
S. iberaensis	Male	B009543	2016	Y	Y	N
S. iberaensis	Male	B009552	2016	Y	Y	N
S. iberaensis	Male	B009560	2016	Y	Y	N
S. iberaensis	Male	B009561	2016	Y	Y	N
S. iberaensis	Male	B009562	2016	Y	Y	N
S. iberaensis	Male	B009583	2016	Y	Y	N
S. iberaensis	Male	B009592	2016	Y	Y	N

S. iberaensis	Female	B009501	2016	Y	Y	N
S. iberaensis	Female	B009505	2016	Y	Y	N
S. iberaensis	Female	B009507	2016	Y	Y	N
S. iberaensis	Female	B009527	2016	Y	Y	N
S. iberaensis	Female	B009542	2016	Y	Y	N
S. iberaensis	Female	B009581	2016	Y	Y	N
S. iberaensis	Female	B009587	2016	Y	Y	N
S. iberaensis	Female	B009590	2016	Y	Y	N
S. iberaensis	Female	B009595	2016	Y	Y	N
S. hypoxantha	Unhatched Egg	UH_Egg_1	2016	N	Y	N
S. hypoxantha	Nestling	B009514	2016	N	Y	Y
S. hypoxantha	Nestling	B009515	2016	N	Y	Y
S. hypoxantha	Nestling	B009516	2016	N	Y	Y
S. hypoxantha	Nestling	B009517	2016	N	Y	Y
S. hypoxantha	Nestling	B009518	2016	N	Y	Y
S. hypoxantha	Nestling	B009519	2016	N	Y	Y
S. hypoxantha	Nestling	B009531	2016	N	Y	Y
S. hypoxantha	Nestling	B009532	2016	N	Y	Y
S. hypoxantha	Nestling	B009537	2016	N	Y	Y
S. hypoxantha	Nestling	B009538	2016	N	Y	Y
S. hypoxantha	Nestling	B009567	2016	N	Y	Y
S. hypoxantha	Nestling	B009569	2016	N	Y	Y
S. hypoxantha	Nestling	B009571	2016	N	Y	Y
S. hypoxantha	Nestling	B009572	2016	N	Y	Y
S. hypoxantha	Nestling	B009575	2016	N	Y	Y
S. hypoxantha	Nestling	B009576	2016	N	Y	Y
S. hypoxantha	Nestling	B009578	2016	N	Y	Y
S. hypoxantha	Nestling	B009579	2016	N	Y	Y
S. hypoxantha	Nestling	B009584	2016	N	Y	Y
S. hypoxantha	Nestling	B009585	2016	N	Y	Y
S. hypoxantha	Male	B009533	2016	N	Y	Y
S. hypoxantha	Male	B009534	2016	N	Y	Y
S. hypoxantha	Male	B009541	2016	N	Y	Y
S. hypoxantha	Male	B009553	2016	N	Y	Y
S. hypoxantha	Male	B009559	2016	N	Y	Y
S. hypoxantha	Male	B009591	2016	N	Y	Y
S. hypoxantha	Female	B009506	2016	N	Y	Y
S. hypoxantha	Female	B009509	2016	N	Y	Y
S. iberaensis	Nestling	B009510	2016	N	Y	Y

S. iberaensis	Nestling	B009511	2016	N	Y	Y
S. iberaensis	Nestling	B009512	2016	N	Y	Y
S. iberaensis	Nestling	B009513	2016	N	Y	Y
S. iberaensis	Nestling	B009535	2016	N	Y	Y
S. iberaensis	Nestling	B009536	2016	N	Y	Y
S. iberaensis	Nestling	B009544	2016	N	Y	Y
S. iberaensis	Nestling	B009545	2016	N	Y	Y
S. iberaensis	Nestling	B009546	2016	N	Y	Y
S. iberaensis	Nestling	B009547	2016	N	Y	Y
S. iberaensis	Nestling	B009548	2016	N	Y	Y
S. iberaensis	Nestling	B009549	2016	N	Y	Y
S. iberaensis	Nestling	B009551	2016	N	Y	Y
S. iberaensis	Nestling	B009588	2016	N	Y	Y
S. iberaensis	Nestling	B009589	2016	N	Y	Y
S. iberaensis	Nestling	B009593	2016	N	Y	Y
S. iberaensis	Nestling	B009594	2016	N	Y	Y
S. iberaensis	Nestling	B009596	2016	N	Y	Y
S. iberaensis	Nestling	B009597	2016	N	Y	Y
S. iberaensis	Male	B009539	2016	N	Y	Y
S. iberaensis	Male	B009550	2016	N	Y	Y
S. iberaensis	Male	B009563	2016	N	Y	Y
S. iberaensis	Male	B009564	2016	N	Y	Y
S. iberaensis	Male	B009565	2016	N	Y	Y
S. iberaensis	Male	B009566	2016	N	Y	Y
S. iberaensis	Male	B009582	2016	N	Y	Y
S. iberaensis	Female	B009586	2016	N	Y	Y
S. hypoxantha	Unhatched Egg	UH_Egg_2	2018	N	Y	Y
S. hypoxantha	Nestling	B009063	2018	N	Y	Y
S. hypoxantha	Nestling	B009064	2018	N	Y	Y
S. hypoxantha	Nestling	B009079	2018	N	Y	Y
S. hypoxantha	Nestling	B009080	2018	N	Y	Y
S. hypoxantha	Nestling	B009108	2018	N	Y	Y
S. hypoxantha	Nestling	B009109	2018	N	Y	Y
S. hypoxantha	Nestling	B009110	2018	N	Y	Y
S. hypoxantha	Nestling	B009111	2018	N	Y	Y
S. hypoxantha	Nestling	B009112	2018	N	Y	Y
S. hypoxantha	Nestling	B009607	2018	N	Y	Y
S. hypoxantha	Nestling	B009608	2018	N	Y	Y
S. hypoxantha	Nestling	B009609	2018	N	Y	Y

S. hypoxantha	Nestling	B009610	2018	N	Y	Y
S. hypoxantha	Nestling	B009613	2018	N	Y	Y
S. hypoxantha	Nestling	B009614	2018	N	Y	Y
S. hypoxantha	Nestling	B009622	2018	N	Y	Y
S. hypoxantha	Nestling	B009623	2018	N	Y	Y
S. hypoxantha	Nestling	B009625	2018	N	Y	Y
S. hypoxantha	Nestling	B009626	2018	N	Y	Y
S. hypoxantha	Nestling	Nestling_1	2018	N	Y	Y
S. hypoxantha	Male	B009027	2018	N	Y	Y
S. hypoxantha	Male	B009028	2018	N	Y	Y
S. hypoxantha	Male	B009030	2018	N	Y	Y
S. hypoxantha	Male	B009031	2018	N	Y	Y
S. hypoxantha	Male	B009032	2018	N	Y	Y
S. hypoxantha	Male	B009033	2018	N	Y	Y
S. hypoxantha	Male	B009035	2018	N	Y	Y
S. hypoxantha	Male	B009055	2018	N	Y	Y
S. hypoxantha	Male	B009056	2018	N	Y	Y
S. hypoxantha	Male	B009057	2018	N	Y	Y
S. hypoxantha	Male	B009059	2018	N	Y	Y
S. hypoxantha	Male	B009060	2018	N	Y	Y
S. hypoxantha	Male	B009067	2018	N	Y	Y
S. hypoxantha	Male	B009077	2018	N	Y	Y
S. hypoxantha	Male	B009078	2018	N	Y	N
S. hypoxantha	Male	B009107	2018	N	Y	Y
S. hypoxantha	Male	B009116	2018	N	Y	Y
S. hypoxantha	Male	B009117	2018	N	Y	Y
S. hypoxantha	Male	B009118	2018	N	Y	Y
S. hypoxantha	Male	B009119	2018	N	Y	Y
S. hypoxantha	Male	B009122	2018	N	Y	Y
S. hypoxantha	Male	B009133	2018	N	Y	N
S. hypoxantha	Male	B009134	2018	N	Y	Y
S. hypoxantha	Male	B009601	2018	N	Y	Y
S. hypoxantha	Male	B009612	2018	N	Y	Y
S. hypoxantha	Male	B009615	2018	N	Y	Y
S. hypoxantha	Female	B009050	2018	N	Y	Y
S. hypoxantha	Female	B009051	2018	N	Y	Y
S. hypoxantha	Female	B009053	2018	N	Y	Y
S. hypoxantha	Female	B009066	2018	N	Y	Y
S. hypoxantha	Female	B009072	2018	N	Y	Y

S. hypoxantha	Female	B009073	2018	N	Y	Y
S. hypoxantha	Female	B009081	2018	N	Y	Y
S. hypoxantha	Female	B009520	2018	N	Y	Y
S. hypoxantha	Female	B009606	2018	N	Y	Y
S. hypoxantha	Female	B009616	2018	N	Y	Y
S. hypoxantha	Female	B009617	2018	N	Y	Y
S. hypoxantha	Female	B009618	2018	N	Y	Y
S. hypoxantha	Female	B009621	2018	N	Y	Y
S. iberaensis	Unhatched Egg	UH_Egg_3	2018	N	Y	N
S. iberaensis	Nestling	B009036	2018	N	Y	Y
S. iberaensis	Nestling	B009037	2018	N	Y	Y
S. iberaensis	Nestling	B009038	2018	N	Y	Y
S. iberaensis	Nestling	B009061	2018	N	Y	Y
S. iberaensis	Nestling	B009062	2018	N	Y	Y
S. iberaensis	Nestling	B009106	2018	N	Y	Y
S. iberaensis	Nestling	B009123	2018	N	Y	Y
S. iberaensis	Nestling	B009124	2018	N	Y	Y
S. iberaensis	Nestling	B009130	2018	N	Y	Y
S. iberaensis	Nestling	B009131	2018	N	Y	Y
S. iberaensis	Nestling	B009603	2018	N	Y	Y
S. iberaensis	Nestling	B009604	2018	N	Y	Y
S. iberaensis	Nestling	B009605	2018	N	Y	Y
S. iberaensis	Nestling	B009627	2018	N	Y	Y
S. iberaensis	Nestling	B009628	2018	N	Y	Y
S. iberaensis	Nestling	B009634	2018	N	Y	Y
S. iberaensis	Nestling	B009635	2018	N	Y	Y
S. iberaensis	Nestling	Nestling_2	2018	N	Y	Y
S. iberaensis	Male	B009029	2018	N	Y	Y
S. iberaensis	Male	B009040	2018	N	Y	Y
S. iberaensis	Male	B009041	2018	N	Y	Y
S. iberaensis	Male	B009042	2018	N	Y	Y
S. iberaensis	Male	B009043	2018	N	Y	Y
S. iberaensis	Male	B009046	2018	N	Y	Y
S. iberaensis	Male	B009047	2018	N	Y	Y
S. iberaensis	Male	B009049	2018	N	Y	Y
S. iberaensis	Male	B009054	2018	N	Y	Y
S. iberaensis	Male	B009058	2018	N	Y	Y
S. iberaensis	Male	B009074	2018	N	Y	Y
S. iberaensis	Male	B009076	2018	N	Y	Y

S. iberaensis	Male	B009113	2018	N	Y	Y
S. iberaensis	Male	B009114	2018	N	Y	Y
S. iberaensis	Male	B009115	2018	N	Y	Y
S. iberaensis	Male	B009121	2018	N	Y	Y
S. iberaensis	Male	B009128	2018	N	Y	Y
S. iberaensis	Male	B009132	2018	N	Y	Y
S. iberaensis	Male	B009139	2018	N	Y	Y
S. iberaensis	Male	B009141	2018	N	Y	Y
S. iberaensis	Male	B009602	2018	N	Y	Y
S. iberaensis	Male	B009620	2018	N	Y	Y
S. iberaensis	Male	B009630	2018	N	Y	Y
S. iberaensis	Female	B009052	2018	N	Y	Y
S. iberaensis	Female	B009083	2018	N	Y	Y
S. iberaensis	Female	B009085	2018	N	Y	Y
S. iberaensis	Female	B009086	2018	N	Y	Y
S. iberaensis	Female	B009120	2018	N	Y	Y
S. iberaensis	Female	B009127	2018	N	Y	Y
S. iberaensis	Female	B009135	2018	N	Y	Y
S. iberaensis	Female	B009140	2018	N	Y	Y
S. iberaensis	Female	B009619	2018	N	Y	Y
S. iberaensis	Female	B009629	2018	N	Y	Y
S. iberaensis	Female	B009633	2018	N	Y	Y

Table S4.2. Sequencing statistics for the whole-genome data.

Band No.	Species	Raw Reads (M)	Expected Coverage[1]	Retained Reads (M)	% Discarded (QC)	Expected Coverage[2]	% Alignment	Depth of Coverage[3]	% Missing in VCF
B009522	*S. hypoxantha*	45.3	5.7	42.9	5.2	5.7	98.8	5.5	3.8
B009524	*S. hypoxantha*	46.2	5.8	42.8	7.4	5.6	98.7	5.6	3.7
B009525	*S. hypoxantha*	57.1	7.2	51.8	9.2	6.8	98.7	6.9	2.0
B009526	*S. hypoxantha*	57.2	7.2	50.8	11.2	6.7	98.9	6.9	1.8
B009530	*S. hypoxantha*	42.3	5.3	40.0	5.3	5.3	98.7	5.1	4.9
B009554	*S. hypoxantha*	38.0	4.8	34.8	8.4	4.6	98.4	4.6	7.5
B009556	*S. hypoxantha*	45.0	5.7	40.9	9.1	5.4	98.5	5.4	4.6
B009574	*S. hypoxantha*	59.1	7.4	54.2	8.3	7.1	98.5	7.1	2.2
B009503	*S. hypoxantha*	38.5	4.8	35.8	6.9	4.7	98.5	4.6	7.2
B009523	*S. hypoxantha*	45.2	5.7	41.3	8.7	5.4	98.6	5.4	4.1
B009529	*S. hypoxantha*	53.2	6.7	50.3	5.5	6.6	98.6	6.4	2.6
B009555	*S. hypoxantha*	43.2	5.4	39.0	9.8	5.1	98.3	5.1	5.4
B009557	*S. hypoxantha*	41.4	5.2	38.4	7.3	5.1	98.5	5.0	5.9
B009570	*S. hypoxantha*	47.5	6.0	44.3	6.8	5.8	98.3	5.7	4.4
B009573	*S. hypoxantha*	48.8	6.1	46.2	5.4	6.1	98.5	5.9	4.1
B009580	*S. hypoxantha*	55.8	7.0	52.5	5.9	6.9	98.4	6.7	2.9
B009502	*S. iberaensis*	43.7	5.5	41.4	5.1	5.5	98.7	5.3	4.6
B009504	*S. iberaensis*	47.2	5.9	44.5	5.7	5.9	98.7	5.7	3.6
B009508	*S. iberaensis*	42.1	5.3	40.2	4.6	5.3	98.7	5.1	5.0
B009528	*S. iberaensis*	37.5	4.7	35.6	5.1	4.7	98.7	4.6	6.7
B009540	*S. iberaensis*	32.0	4.0	30.2	5.5	4.0	98.8	3.9	9.5
B009543	*S. iberaensis*	48.0	6.0	45.5	5.3	6.0	98.6	5.8	3.4
B009552	*S. iberaensis*	39.0	4.9	36.3	6.9	4.8	98.5	4.7	6.8
B009560	*S. iberaensis*	40.0	5.0	37.8	5.5	5.0	98.6	4.9	6.6
B009561	*S. iberaensis*	49.9	6.3	47.3	5.2	6.2	98.5	6.0	4.0
B009562	*S. iberaensis*	31.3	3.9	29.3	6.3	3.9	98.3	3.8	12.3
B009583	*S. iberaensis*	57.5	7.2	53.5	6.9	7.0	98.5	6.9	2.4
B009592	*S. iberaensis*	46.1	5.8	41.1	10.7	5.4	98.6	5.5	4.2
B009501	*S. iberaensis*	43.1	5.4	40.4	6.3	5.3	98.8	5.2	4.5
B009505	*S. iberaensis*	48.7	6.1	45.6	6.4	6.0	98.7	5.9	3.3
B009507	*S. iberaensis*	40.3	5.1	37.7	6.5	5.0	98.6	4.9	5.8
B009527	*S. iberaensis*	46.6	5.9	44.2	5.3	5.8	98.5	5.6	4.0
B009542	*S. iberaensis*	44.1	5.6	41.6	5.8	5.5	98.6	5.3	4.5
B009581	*S. iberaensis*	41.9	5.3	39.5	5.6	5.2	98.5	5.1	6.1

B009587	*S. iberaensis*	49.8	6.3	45.1	9.5	5.9	98.5	6.0	3.6
B009590	*S. iberaensis*	44.1	5.6	38.3	13.2	5.0	98.4	5.2	5.2
B009595	*S. iberaensis*	85.2	10.7	80.0	6.1	10.5	98.4	10.3	0.8

[1] Computed as the number of raw 151 bp reads divided by a genome size of ~1.2 Gbp.

[2] Computed as the number of reads retained after filtering multiplied by their average size, and then divided by a genome size of ~1.2 Gbp.

[3] Calculated from the bam files after alignment.

Table S4.3. Position of highly divergent SNPs ($F_{ST} > 0.85$) relative to identified genes.
The ~37 kb of sequence between *OCA2* and *HERC2* produced a mean PhastCons conservation score of 0.048 ± 0.015 SD. However, these values ranged from 0 to up to 0.997, with defined peaks and 215 positions showing a conservation score of 0.9 or higher. It is likely that these regions, which are conserved among distantly related species, contain cis-regulatory elements that are necessary to control the expression of *OCA2*. The highly divergent SNPs were located between 6,496 and 8,711 kb upstream of *OCA2* and from 30,506 kb downstream of *HERC2* to an intron within that gene. There was one highly differentiated SNP 17,815 kb upstream of *TYRP1*.

Scaffold	Position of highly divergent SNP	F_{ST} value	Location relative to identified genes
430	11027082	0.852	Upstream of *OCA2*, Downstream of *HERC2*
430	11027706	0.857	Upstream of *OCA2*, Downstream of *HERC2*
430	11028530	0.915	Upstream of *OCA2*, Downstream of *HERC2*
430	11028725	0.874	Upstream of *OCA2*, Downstream of *HERC2*
430	11028788	0.887	Upstream of *OCA2*, Downstream of *HERC2*
430	11028845	0.884	Upstream of *OCA2*, Downstream of *HERC2*
430	11028855	0.941	Upstream of *OCA2*, Downstream of *HERC2*
430	11028917	0.884	Upstream of *OCA2*, Downstream of *HERC2*
430	11028925	0.882	Upstream of *OCA2*, Downstream of *HERC2*
430	11028946	0.858	Upstream of *OCA2*, Downstream of *HERC2*
430	11028951	0.858	Upstream of *OCA2*, Downstream of *HERC2*
430	11028961	0.858	Upstream of *OCA2*, Downstream of *HERC2*
430	11029143	0.855	Upstream of *OCA2*, Downstream of *HERC2*
430	11029157	0.883	Upstream of *OCA2*, Downstream of *HERC2*
430	11029206	0.916	Upstream of *OCA2*, Downstream of *HERC2*
430	11029270	0.885	Upstream of *OCA2*, Downstream of *HERC2*
430	11029304	0.885	Upstream of *OCA2*, Downstream of *HERC2*
430	11030000	0.862	Upstream of *OCA2*, Downstream of *HERC2*
430	11030054	0.884	Upstream of *OCA2*, Downstream of *HERC2*
430	11030110	0.884	Upstream of *OCA2*, Downstream of *HERC2*

430	11043140	0.879	Upstream of *OCA2*, Downstream of *HERC2*
430	11044894	0.878	Upstream of *OCA2*, Downstream of *HERC2*
430	11054240	0.933	Upstream of *OCA2*, Downstream of *HERC2*
430	11054328	0.885	Upstream of *OCA2*, Downstream of *HERC2*
430	11058206	0.858	Upstream of *OCA2*, Downstream of *HERC2*
430	11058495	0.875	Exon of *HERC2*
430	11058847	0.880	Intron of *HERC2*
430	11059368	0.873	Intron of *HERC2*
430	11060287	0.874	Intron of *HERC2*
430	11060381	0.906	Intron of *HERC2*
430	11060579	0.850	Intron of *HERC2*
430	11064841	0.850	Intron of *HERC2*
430	11064842	0.850	Intron of *HERC2*
430	11067576	0.879	Intron of *HERC2*
430	11067689	0.852	Intron of *HERC2*
430	11067987	0.883	Intron of *HERC2*
257	21661286	0.868	Intron of *TYRP1*
257	21661528	0.884	Intron of *TYRP1*
257	21661610	0.862	Intron of *TYRP1*
257	21687503	0.864	Upstream of *TYRP1*
762	1684725	0.878	Intron of *GTP2*
762	1685092	0.862	Intron of *GTP2*

Table S4.4. Parentage assignments from the ddRAD data.

Nestling Band No.	Social Father Band No.	Social Father Species	Candidate Father Band No.	Candidate Father Species	No. Mismatching Loci (out of 281)	Extra-pair?
B009036	B009128	*S. iberaensis*	B009563	*S. iberaensis*	7 (2.5%)	Y
B009037	B009121	*S. iberaensis*	B009058	*S. iberaensis*	0 (0%)	Y
B009038	B009121	*S. iberaensis*	B009058	*S. iberaensis*	2 (0.7%)	Y
B009061	B009046	*S. iberaensis*	Unassigned	Unassigned	23 (8.2%)	Y
B009062	B009046	*S. iberaensis*	B009046	*S. iberaensis*	5 (1.8%)	N
B009063	B009030	*S. hypoxantha*	B009030	*S. hypoxantha*	8 (2.8%)	N
B009064	B009030	*S. hypoxantha*	B009030	*S. hypoxantha*	2 (0.7%)	N
B009079	B009067	*S. hypoxantha*	B009030	*S. hypoxantha*	5 (1.8%)	Y
B009080	B009067	*S. hypoxantha*	B009030	*S. hypoxantha*	3 (1.1%)	Y
B009106	B009141	*S. iberaensis*	B009141	*S. iberaensis*	2 (0.7%)	N

B009108	B009117	*S. hypoxantha*	B009116	*S. hypoxantha*	7 (2.5%)	Y
B009109	B009117	*S. hypoxantha*	B009116	*S. hypoxantha*	4 (1.4%)	Y
B009110	B009117	*S. hypoxantha*	B009117	*S. hypoxantha*	2 (0.7%)	N
B009111	B009116	*S. hypoxantha*	B009116	*S. hypoxantha*	3 (1.1%)	N
B009112	B009116	*S. hypoxantha*	B009116	*S. hypoxantha*	5 (1.8%)	N
B009123	B009115	*S. iberaensis*	B009115	*S. iberaensis*	3 (1.1%)	N
B009124	B009115	*S. iberaensis*	B009115	*S. iberaensis*	3 (1.1%)	N
B009130	B009113	*S. iberaensis*	B009113	*S. iberaensis*	2 (0.7%)	N
B009131	B009113	*S. iberaensis*	B009528	*S. iberaensis*	1 (0.4%)	Y
B009510	B009508	*S. iberaensis*	B009528	*S. iberaensis*	6 (2.1%)	Y
B009511	B009508	*S. iberaensis*	B009528	*S. iberaensis*	6 (2.1%)	Y
B009512	B009504	*S. iberaensis*	Unassigned	Unassigned	22 (7.8%)	Y
B009513	B009504	*S. iberaensis*	B009504	*S. iberaensis*	0 (0%)	N
B009514	B009522	*S. hypoxantha*	B009522	*S. hypoxantha*	5 (1.8%)	N
B009515	B009522	*S. hypoxantha*	B009522	*S. hypoxantha*	3 (1.1%)	N
B009516	Unbanded	*S. hypoxantha*	Unassigned	Unassigned	23 (8.2%)	Unknown
B009517	Unbanded	*S. hypoxantha*	Unassigned	Unassigned	24 (8.5%)	Unknown
B009518	Unbanded	*S. hypoxantha*	Unassigned	Unassigned	22 (7.8%)	Unknown
B009519	Unbanded	*S. hypoxantha*	Unassigned	Unassigned	20 (7.1%)	Unknown
B009531	Unbanded	*S. hypoxantha*	Unassigned	Unassigned	20 (7.1%)	Unknown
B009532	Unbanded	*S. hypoxantha*	Unassigned	Unassigned	22 (7.8%)	Unknown
B009535	B009528	*S. iberaensis*	B009528	*S. iberaensis*	2 (0.7%)	N
B009536	B009528	*S. iberaensis*	B009508	*S. iberaensis*	2 (0.7%)	Y
B009537	B009524	*S. hypoxantha*	B009524	*S. hypoxantha*	5 (1.8%)	N
B009538	B009524	*S. hypoxantha*	B009524	*S. hypoxantha*	0 (0%)	N
B009544	B009539	*S. iberaensis*	B009583	*S. iberaensis*	2 (0.7%)	Y
B009545	B009539	*S. iberaensis*	Unassigned	Unassigned	18 (6.4%)	Y
B009546	B009540	*S. iberaensis*	B009539	*S. iberaensis*	5 (1.8%)	Y
B009547	B009540	*S. iberaensis*	B009539	*S. iberaensis*	4 (1.4%)	Y
B009548	B009550	*S. iberaensis*	B009550	*S. iberaensis*	2 (0.7%)	N
B009549	B009550	*S. iberaensis*	B009550	*S. iberaensis*	4 (1.4%)	N
B009551	Unbanded	*S. iberaensis*	Unassigned	Unassigned	17 (6.0%)	Unknown
B009567	B009556	*S. hypoxantha*	Unassigned	Unassigned	16 (5.7%)	Y
B009569	B009556	*S. hypoxantha*	B009556	*S. hypoxantha*	5 (1.8%)	N
B009571	B009526	*S. hypoxantha*	B009553	*S. hypoxantha*	7 (2.5%)	Y
B009572	B009526	*S. hypoxantha*	B009553	*S. hypoxantha*	6 (2.1%)	Y
B009575	B009522	*S. hypoxantha*	Unassigned	Unassigned	19 (6.8%)	Y
B009576	B009522	*S. hypoxantha*	B009522	*S. hypoxantha*	3 (1.1%)	N
B009578	B009530	*S. hypoxantha*	B009530	*S. hypoxantha*	5 (1.8%)	N

B009579	Unbanded	*S. hypoxantha*	B009541	*S. hypoxantha*	3 (1.1%)	Unknown
B009584	B009554	*S. hypoxantha*	B009554	*S. hypoxantha*	2 (0.7%)	N
B009585	B009554	*S. hypoxantha*	B009522	*S. hypoxantha*	3 (1.1%)	Y
B009588	B009562	*S. iberaensis*	Unassigned	Unassigned	26 (9.3%)	Y
B009589	B009562	*S. iberaensis*	Unassigned	Unassigned	21 (7.5%)	Y
B009593	B009552	*S. iberaensis*	Unassigned	Unassigned	20 (7.1%)	Y
B009594	B009552	*S. iberaensis*	Unassigned	Unassigned	17 (6.0%)	Y
B009596	B009583	*S. iberaensis*	B009583	*S. iberaensis*	0 (0%)	N
B009597	B009583	*S. iberaensis*	B009583	*S. iberaensis*	0 (0%)	N
B009603	B009042	*S. iberaensis*	B009042	*S. iberaensis*	2 (0.7%)	N
B009604	B009042	*S. iberaensis*	Unassigned	Unassigned	27 (9.6%)	Y
B009605	B009042	*S. iberaensis*	Unassigned	Unassigned	24 (8.5%)	Y
B009607	B009055	*S. hypoxantha*	Unassigned	Unassigned	16 (5.7%)	Y
B009608	B009055	*S. hypoxantha*	B009055	*S. hypoxantha*	4 (1.4%)	N
B009609	B009601	*S. hypoxantha*	Unassigned	Unassigned	20 (7.1%)	Y
B009610	B009601	*S. hypoxantha*	Unassigned	Unassigned	20 (7.1%)	Y
B009613	B009612	*S. hypoxantha*	B009612	*S. hypoxantha*	4 (1.4%)	N
B009614	B009612	*S. hypoxantha*	B009612	*S. hypoxantha*	5 (1.8%)	N
B009622	B009060	*S. hypoxantha*	B009060	*S. hypoxantha*	2 (0.7%)	N
B009623	B009060	*S. hypoxantha*	Unassigned	Unassigned	23 (8.2%)	Y
B009625	B009118	*S. hypoxantha*	B009118	*S. hypoxantha*	3 (1.1%)	N
B009626	B009118	*S. hypoxantha*	Unassigned	Unassigned	27 (9.6%)	Y
B009627	B009620	*S. iberaensis*	B009620	*S. iberaensis*	5 (1.8%)	N
B009628	B009620	*S. iberaensis*	B009508	*S. iberaensis*	3 (1.1%)	Y
B009634	Unbanded	*S. iberaensis*	Unassigned	Unassigned	19 (6.8%)	Unknown
B009635	Unbanded	*S. iberaensis*	Unassigned	Unassigned	21 (7.5%)	Unknown
Nestling_1	B009119	*S. hypoxantha*	Unassigned	Unassigned	20 (7.1%)	Y
Nestling_2	B009128	*S. iberaensis*	B009128	*S. iberaensis*	2 (0.7%)	N
UH_Egg_1	Unbanded	*S. hypoxantha*	Unassigned	Unassigned	0 (0%)	Unknown
UH_Egg_2	Unbanded	*S. hypoxantha*	Unassigned	Unassigned	24 (8.5%)	Unknown
UH_Egg_3	Unbanded	*S. iberaensis*	Unassigned	Unassigned	19 (6.8%)	Unknown

Table S4.5. Post-hoc pairwise comparisons (Tukey HSD method) for generalized linear mixed models examining the effect of treatment group on behavioral response intensity (PC1). Territorial males of *S. hypoxantha* (top) and *S. iberaensis* (bottom) were presented with combinations of conspecific (CON), heterospecific (HET), and *S. collaris* (CONTROL) song and plumage. Significant results (adjusted $P < 0.05$) are highlighted in bold. Treatment group had a significant effect on response intensity in both *S. hypoxantha* ($P < 0.0001$, $N = 120$) and *S. iberaensis* ($P < 0.0001$, $N = 120$) males. Asterisks indicate treatments groups that were not significantly different when outliers (observations outside 1.5 * interquartile range; *S. hypoxantha*: 11 outliers, *S. iberaensis*: 11 outliers) were removed.*

S. hypoxantha

Treatment contrast[1]	**Estimate**	**S.E.**	**df**	***P* value**
CONTROL – CON Plumage/CON Song	-3.10	0.28	90.7	**<0.0001**
CONTROL – CON Plumage/HET Song	-1.80	0.27	85.6	**<0.0001**
CONTROL – HET Plumage/CON Song	-2.17	0.28	90.7	**<0.0001**
CONTROL – HET Plumage/HET Song	-0.60	0.28	89.0	0.2078
CON Plumage/CON Song – CON Plumage/HET Song	1.30	0.27	85.0	**0.0001**
CON Plumage/CON Song – HET Plumage/CON Song*	0.93	0.27	87.9	**0.0082**
CON Plumage/CON Song – HET Plumage/HET Song	2.51	0.28	89.9	**<0.0001**
CON Plumage/HET Song – HET Plumage/CON Song	-0.37	0.27	85.0	0.6544
CON Plumage/HET Song – HET Plumage/HET Song	1.21	0.27	83.6	**0.0002**
HET Plumage/CON Song – HET Plumage/HET Song	1.57	0.28	89.9	**<0.0001**

S. iberaensis

Treatment contrast[2]	**Estimate**	**S.E.**	**df**	***P* value**
CONTROL – HET Plumage/HET Song	-0.84	0.26	86.6	**0.0136**
CONTROL – HET Plumage/CON Song	-2.57	0.26	87.8	**<0.0001**
CONTROL – CON Plumage/HET Song	-1.99	0.26	91.4	**<0.0001**
CONTROL – CON Plumage/CON Song	-3.47	0.27	92.5	**<0.0001**
CON Plumage/CON Song – CON Plumage/HET Song	-1.48	0.26	90.4	**<0.0001**
CON Plumage/CON Song – HET Plumage/CON Song*	-0.90	0.26	86.8	**0.0068**
CON Plumage/CON Song – HET Plumage/HET Song	-2.62	0.26	90.4	**<0.0001**
CON Plumage/HET Song – HET Plumage/CON Song	0.58	0.26	87.9	0.1745
CON Plumage/HET Song – HET Plumage/HET Song	-1.15	0.26	89.1	**0.0003**
HET Plumage/CON Song – HET Plumage/HET Song	-1.73	0.26	86.5	**<0.0001**

[1]Model included male ID (SD = 0.90; 95% CI of SD = 0.64-1.21) as a random effect.

[2]Model included male ID (SD = 0.72; 95% CI of SD = 0.50-1.02) and female presence (SD = 0.46; 95% CI of SD = 0.09-1.49) as random effects.

*Although we did not detect a significant difference between the CON Plumage/CON Song treatment and the HET Plumage/CON Song treatment in both species, suggesting that song may play a more important role than plumage in territorial interactions, this result was only obtained after excluding outliers and further research is necessary to evaluate the relative importance of these two divergent traits.

APPENDIX A5

Supplemental Figures and Tables for Chapter 5

SUPPLEMENTAL FIGURES

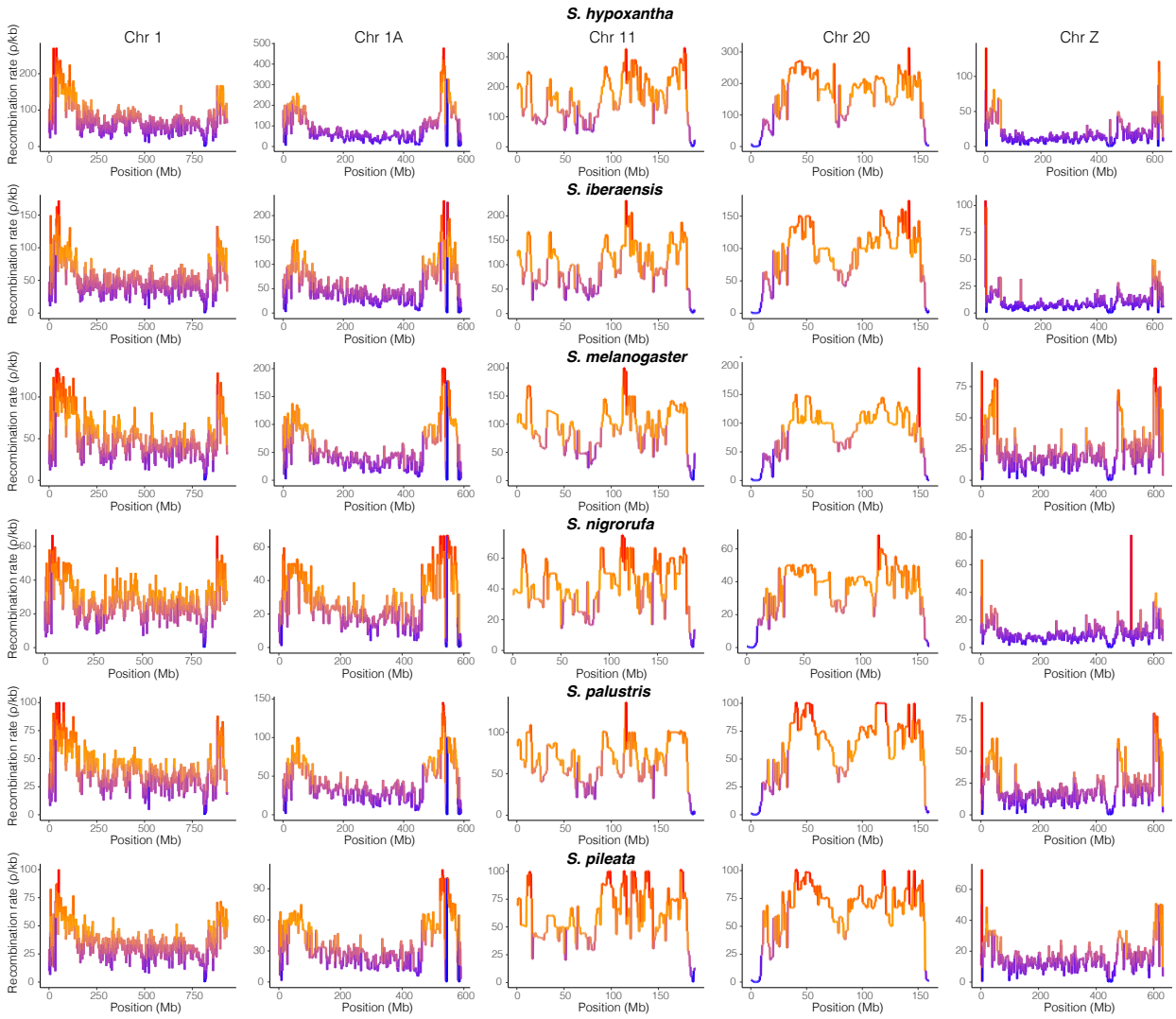


Figure S5.1. Fine-scale recombination maps across chromosomes 1, 1A, 11, 20, and Z for six species of southern capuchino seedeater: (top to bottom) *S. hypoxantha*, *S. iberaensis*, *S. melanogaster*, *S. nigroufa*, *S. palustris*, and *S. pileata*. Recombination estimates were averaged over 100-kb windows and are reported in ρ/bp, where ρ is the population-scaled recombination rate ($4N_er$) and r is the per-generation recombination rate (*i.e.,* the probability of a recombination event occurring during meiosis).

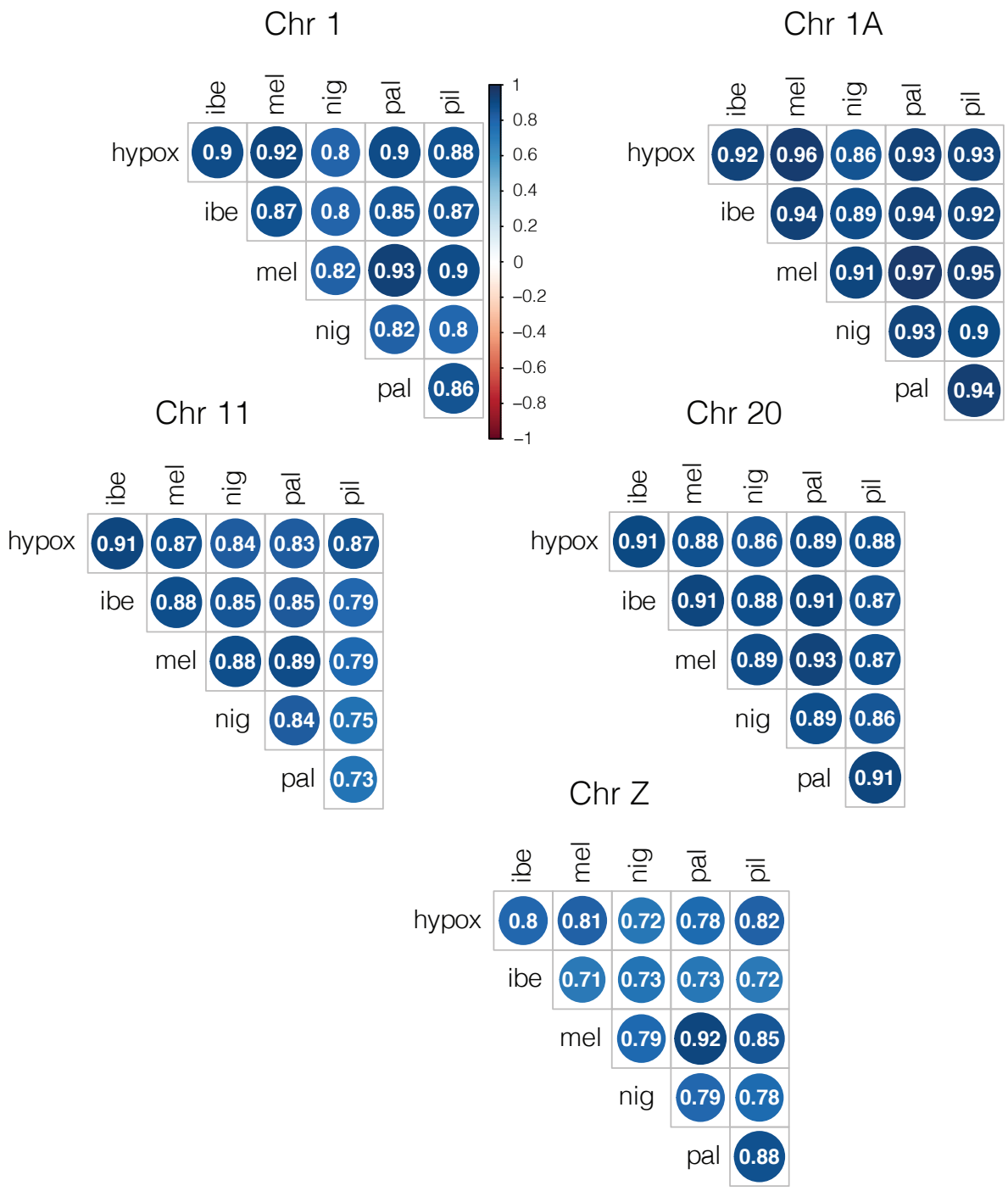


Figure S5.2. Correlation matrices showing the conservation of recombination rates across chromosomes 1, 1A, 11, 20, and Z between six southern capuchino species: *S. hypoxantha* (hypox), *S. iberaensis* (ibe), *S. melanogaster* (mel), *S. nigrorufa* (nig), *S. palustris* (pal), and *S. pileata* (pil). Other details as in Figure S5.1.

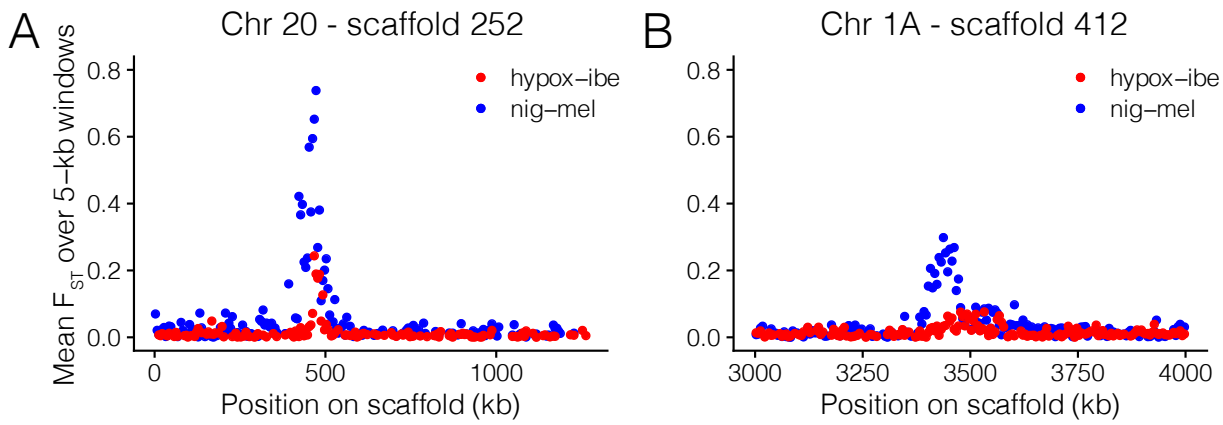


Figure S5.3. Example plots showing that divergence peaks were specific to each focal comparison despite conserved recombination landscapes among capuchino species. Plots of mean F_{ST} between *S. hypoxantha* (hypox) vs. *S. iberaensis* (ibe) and *S. nigrorufa* (nig) vs. *S. melanogaster* (mel) across **(A)** scaffold 252 on chromosome 20 and **(B)** scaffold 412 on chromosome 1A. While the recombination landscapes of these four capuchino species showed similar patterns across both scaffolds, peaks of divergence were only present in the *S. nigrorufa* vs. *S. melanogaster* comparison. Divergence peaks were defined as regions with an F_{ST} value > 0.2 that contained at least one individual SNP with an F_{ST} value > 0.85 [as in Campagna et al. (2017)].

SUPPLEMENTAL TABLES

Table S5.1. Details on samples included in the study.

Species	Catalogue/band number[a]	Locality	Lat	Lon	Sex
S. pileata	KUNHM 3664	San Rafael National Park, Itapuá, Paraguay	-26.52	-55.80	M
S. pileata	KUNHM 3687	San Rafael National Park, Itapuá, Paraguay	-26.52	-55.80	M
S. pileata	KUNHM 3691	San Rafael National Park, Itapuá, Paraguay	-26.52	-55.80	M
S. pileata	KUNHM 3699	San Rafael National Park, Itapuá, Paraguay	-26.52	-55.80	M
S. pileata	MACN 6537	Captive; Argentina	-	-	M
S. pileata	MZUSP 77832	Santa Gertrudes, São Paulo, Brazil	-22.46	-47.53	M
S. pileata	MCP 3627	Itararé, São Paulo, Brazil	-24.11	-49.33	M
S. pileata	MCP 3628	Itararé, São Paulo, Brazil	-24.11	-49.33	M
S. pileata	MCP 4222	Chapadão do Céu, Goiás, Brazil	-18.40	-52.67	M
S. pileata	MCP 4850	Broa, São Paulo, Brazil	-22.20	-47.87	M
S. pileata	MCP 4851	Dourado, São Paulo, Brazil	-22.12	-48.34	M
S. pileata	MCP 4852	Dourado, São Paulo, Brazil	-22.12	-48.34	M
S. palustris	MACN 3118	Iberá, Corrientes, Argentina	-27.86	-56.7	M
S. palustris	MZUSP 94877	Captive; Brazil	-	-	M
S. palustris	MACN 5173	Iberá, Corrientes, Argentina	-27.86	-56.7	M
S. palustris	MACN 5175	Iberá, Corrientes, Argentina	-27.86	-56.7	M
S. palustris	MACN 5178	Iberá, Corrientes, Argentina	-27.86	-56.7	M
S. palustris	MACN 5179	Iberá, Corrientes, Argentina	-27.86	-56.7	M
S. palustris	MACN 5168	Iberá, Corrientes, Argentina	-27.86	-56.7	M
S. palustris	MACN 5240	Iberá, Corrientes, Argentina	-27.86	-56.7	M
S. palustris	MCP NN1	Captive; Brazil	-	-	M
S. palustris	MCP NN2	Captive; Brazil	-	-	M
S. palustris	MACN 3117	Gualeguaychú, Entre Ríos, Argentina	-33.01	-58.52	M
S. palustris	MACN 3372	Gualeguaychú, Entre Ríos, Argentina	-33.01	-58.52	M
S. melanogaster	MCP 2312	Bom Jesus, Rio Grande do Sul, Brazil	-28.66	-50.44	M
S. melanogaster	MCP 2318	Bom Jesus, Rio Grande do Sul, Brazil	-28.66	-50.44	M
S. melanogaster	MCP 2073	Bom Jesus, Rio Grande do Sul, Brazil	-28.66	-50.44	F
S. melanogaster	MCP 2074	Bom Jesus, Rio Grande do Sul, Brazil	-28.66	-50.44	M
S. melanogaster	MCP 2076	Bom Jesus, Rio Grande do Sul, Brazil	-28.66	-50.44	M
S. melanogaster	MCP 2078	Bom Jesus, Rio Grande do Sul, Brazil	-28.66	-50.44	M
S. melanogaster	MCP 2296	Bañado Água Branca, Bom Jesus, Rio Grande do Sul, Brazil	-28.60	-50.41	M

S. melanogaster	MCP 2298	Bañado Água Branca, Bom Jesus, Rio Grande do Sul, Brazil	-28.60	-50.41	M
S. melanogaster	MCP 2306	Bañado Rio Santana, Bom Jesus, Rio Grande do Sul Brazil	-28.48	-50.72	M
S. melanogaster	MCP 2311	Bañado Água Branca, Bom Jesus, Rio Grande do Sul, Brazil	-28.60	-50.41	M
S. melanogaster	MCP 2315	Coxilha Rica, Lages Santa Catarina, Brazil	-28.31	-50.28	M
S. melanogaster	MCP 2075	Bom Jesus, Rio Grande do Sul, Brazil	-28.66	-50.44	M
S. nigrorufa	MZUSP 98660	Vila Bela Da Santíssima Trindade, Mato Grosso, Brazil	-15.05	-59.92	M
S. nigrorufa	MZUSP 98639	Vila Bela Da Santíssima Trindade, Mato Grosso, Brazil	-15.05	-59.92	M
S. nigrorufa	MZUSP 98590	Vila Bela Da Santíssima Trindade, Mato Grosso, Brazil	-15.05	-59.92	M
S. nigrorufa	MZUSP 98661	Vila Bela Da Santíssima Trindade, Mato Grosso, Brazil	-15.05	-59.92	M
S. nigrorufa	MZUSP 98662	Vila Bela Da Santíssima Trindade, Mato Grosso, Brazil	-15.05	-59.92	M
S. nigrorufa	MZUSP 98637	Vila Bela Da Santíssima Trindade, Mato Grosso, Brazil	-15.05	-59.92	M
S. nigrorufa	MCP 4719	Vila Bela Da Santíssima Trindade, Mato Grosso, Brazil	-15.05	-59.92	M
S. nigrorufa	MCP 4720	Vila Bela Da Santíssima Trindade, Mato Grosso, Brazil	-15.05	-59.92	M
S. nigrorufa	MCP 4723	Vila Bela Da Santíssima Trindade, Mato Grosso, Brazil	-15.05	-59.92	M
S. nigrorufa	MCP 4724	Vila Bela Da Santíssima Trindade, Mato Grosso, Brazil	-15.05	-59.92	M
S. nigrorufa	MCP 4725	Vila Bela Da Santíssima Trindade, Mato Grosso, Brazil	-15.05	-59.92	M
S. nigrorufa	MCP 4727	Vila Bela Da Santíssima Trindade, Mato Grosso, Brazil	-15.05	-59.92	M
S. hypoxantha	MACN 3103	El Bagual, Formosa, Argentina	-26.17	-58.93	M
S. hypoxantha	MACN 3105	El Bagual, Formosa, Argentina	-26.17	-58.93	M
S. hypoxantha	MACN 4970	El Bagual, Formosa, Argentina	-26.17	-58.93	M
S. hypoxantha	MACN 4971	El Bagual, Formosa, Argentina	-26.17	-58.93	M
S. hypoxantha	MACN 4972	El Bagual, Formosa, Argentina	-26.17	-58.93	M
S. hypoxantha	MACN 3098	El Bagual, Formosa, Argentina	-26.17	-58.93	M
S. hypoxantha	MACN 4975	El Bagual, Formosa, Argentina	-26.17	-58.93	M
S. hypoxantha	MACN 3254	El Bagual, Formosa, Argentina	-26.17	-58.93	M

S. hypoxantha	MACN 5177	Iberá, Corrientes, Argentina	-27.86	-56.7	M
S. hypoxantha	MACN 5234	Iberá, Corrientes, Argentina	-27.86	-56.7	M
S. hypoxantha	MACN 3272	Estero Catalina, Formosa, Argentina	-25.11	-58.15	M
S. hypoxantha	MACN 3258	El Bagual, Formosa, Argentina	-26.17	-58.93	M
S. hypoxantha	B009522	Iberá, Corrientes, Argentina	-27.86	-56.7	M
S. hypoxantha	B009524	Iberá, Corrientes, Argentina	-27.86	-56.7	M
S. hypoxantha	B009525	Iberá, Corrientes, Argentina	-27.86	-56.7	M
S. hypoxantha	B009526	Iberá, Corrientes, Argentina	-27.86	-56.7	M
S. hypoxantha	B009530	Iberá, Corrientes, Argentina	-27.86	-56.7	M
S. hypoxantha	B009554	Iberá, Corrientes, Argentina	-27.86	-56.7	M
S. hypoxantha	B009556	Iberá, Corrientes, Argentina	-27.86	-56.7	M
S. hypoxantha	B009574	Iberá, Corrientes, Argentina	-27.86	-56.7	M
S. hypoxantha	B009503	Iberá, Corrientes, Argentina	-27.86	-56.7	F
S. hypoxantha	B009523	Iberá, Corrientes, Argentina	-27.86	-56.7	F
S. hypoxantha	B009529	Iberá, Corrientes, Argentina	-27.86	-56.7	F
S. hypoxantha	B009555	Iberá, Corrientes, Argentina	-27.86	-56.7	F
S. hypoxantha	B009557	Iberá, Corrientes, Argentina	-27.86	-56.7	F
S. hypoxantha	B009570	Iberá, Corrientes, Argentina	-27.86	-56.7	F
S. hypoxantha	B009573	Iberá, Corrientes, Argentina	-27.86	-56.7	F
S. hypoxantha	B009580	Iberá, Corrientes, Argentina	-27.86	-56.7	F
S. iberaensis	B009502	Iberá, Corrientes, Argentina	-27.86	-56.7	M
S. iberaensis	B009504	Iberá, Corrientes, Argentina	-27.86	-56.7	M
S. iberaensis	B009508	Iberá, Corrientes, Argentina	-27.86	-56.7	M
S. iberaensis	B009528	Iberá, Corrientes, Argentina	-27.86	-56.7	M
S. iberaensis	B009540	Iberá, Corrientes, Argentina	-27.86	-56.7	M
S. iberaensis	B009543	Iberá, Corrientes, Argentina	-27.86	-56.7	M
S. iberaensis	B009552	Iberá, Corrientes, Argentina	-27.86	-56.7	M
S. iberaensis	B009560	Iberá, Corrientes, Argentina	-27.86	-56.7	M
S. iberaensis	B009561	Iberá, Corrientes, Argentina	-27.86	-56.7	M
S. iberaensis	B009562	Iberá, Corrientes, Argentina	-27.86	-56.7	M
S. iberaensis	B009583	Iberá, Corrientes, Argentina	-27.86	-56.7	M
S. iberaensis	B009592	Iberá, Corrientes, Argentina	-27.86	-56.7	M
S. iberaensis	B009501	Iberá, Corrientes, Argentina	-27.86	-56.7	F
S. iberaensis	B009505	Iberá, Corrientes, Argentina	-27.86	-56.7	F
S. iberaensis	B009507	Iberá, Corrientes, Argentina	-27.86	-56.7	F
S. iberaensis	B009527	Iberá, Corrientes, Argentina	-27.86	-56.7	F
S. iberaensis	B009542	Iberá, Corrientes, Argentina	-27.86	-56.7	F
S. iberaensis	B009581	Iberá, Corrientes, Argentina	-27.86	-56.7	F
S. iberaensis	B009587	Iberá, Corrientes, Argentina	-27.86	-56.7	F
S. iberaensis	B009590	Iberá, Corrientes, Argentina	-27.86	-56.7	F
S. iberaensis	B009595	Iberá, Corrientes, Argentina	-27.86	-56.7	F
S. castaneiventris	MZUSP SP27	Captive; Brazil	-	-	M
S. castaneiventris	MZUSP SP33	Captive; Brazil	-	-	M
S. minuta	CUMV 55642	Parque Nacional Laguna de Chacahua, Oaxaca, Mexico	16.02	-97.76	M

S. minuta	CUMV 55724	Parque Nacional Laguna de Chacahua, Oaxaca, Mexico	16.02	-97.76	M
D. indicus	KU (4177) DA-01	China	42	120.93	M
D. indicus	KU (4208) DA-02	China	45.3	127.36	F
D. indicus	KU (4182) DA-03	China	42	120.93	M
D. indicus	IOZ 10675 /DD 09051	Liaoning Prov., China	40.13	124.35	M
D. indicus	IOZ 14114 /FJ0903	Fujian, China	26.13	119.36	Unknown
Mot cinerea cinerea	UWBM 61135 / IUK 280	Krasnodar, Krasnodarskiy Kray, Russia	44.7	38.82	M
M. cinerea cinerea	UWBM 61149 / IUK 293	Krasnodar, Krasnodarskiy Kray, Russia	44.7	38.82	M
M. cinerea cinerea	ZMUK 148951 / JBK25-13.2.13	Liaoning Prov., China	57.57	10.11	Unknown
M. cinerea cinerea	IOZ 10592 / SX 009	Fujian, China	33.74	107.42	Unknown
M. alba alba	DRK0381	Tiumenskay oblast, Russia	71.38	56.04	M
M. alba alba	DRK0382	Tiumenskay oblast, Russia	71.38	56.04	M
M. alba alba	DRK0383	Tiumenskay oblast, Russia	70.72	56.29	M
M. alba alba	DRK0384	Tiumenskay oblast, Russia	70.11	56.20	M
M. alba alba	DRK0385	Tiumenskay oblast, Russia	70.44	56.31	M
M. alba alba	DRK0386	Tiumenskay oblast, Russia	70.44	56.31	M
M. alba alba	DRK0388	Tiumenskay oblast, Russia	70.61	56.42	M
M. alba alba	DRK0389	Tiumenskay oblast, Russia	70.61	56.42	M
M. alba alba	DRK0391	Tiumenskay oblast, Russia	70.86	56.27	M
M. alba alba	DRK0392	Tiumenskay oblast, Russia	70.86	56.27	M
M. alba alba	SGA1665	Altaiskii kray, Russia	53.22	84.68	M
M. alba alba	SGA1678	Altaiskii kray, Russia	52.89	84.75	M
M. alba alba	SGA1680	Altaiskii kray, Russia	52.89	84.75	M
M. alba alba	SGA1687	Altaiskii kray, Russia	53.22	84.51	M
M. alba alba	SGA1690	Altaiskii kray, Russia	53.22	84.51	M
M. alba alba	SGA1694	Altaiskii kray, Russia	52.94	84.59	M
M. alba alba	SGA1695	Altaiskii kray, Russia	52.94	84.59	M
M. alba alba	SGA1700	Altaiskii kray, Russia	52.85	84.48	M
M. alba alba	SGA1703	Altaiskii kray, Russia	52.85	84.48	M
M. alba alba	SGA1717	Altaiskii kray, Russia	52.80	84.92	M
M. alba alba	SGA1724	Altaiskii kray, Russia	52.80	84.84	M
M. alba alba	SGA1733	Altaiskii kray, Russia	52.73	84.96	M
M. alba alba	SGA1736	Altaiskii kray, Russia	52.73	84.96	M
M. alba alba	SGA1737	Altaiskii kray, Russia	52.73	84.96	M
M. alba alba	SGA1739	Altaiskii kray, Russia	52.73	84.96	M
M. alba alba	SGA1774	Altaiskii kray, Russia	52.65	85.17	M
M. alba alba	SGA1786	Altaiskii kray, Russia	52.38	85.67	M
M. alba alba	SGA1798	Altaiskii kray, Russia	52.28	85.44	M

M. alba personata	SGA1738	Altaiskii kray, Russia	52.73	84.96	M
M. alba personata	SGA1765	Altaiskii kray, Russia	52.60	85.06	M
M. alba personata	SGA1768	Altaiskii kray, Russia	52.60	85.06	M
M. alba personata	SGA1778	Altaiskii kray, Russia	52.65	85.17	M
M. alba personata	SGA1781	Altaiskii kray, Russia	52.65	85.17	M
M. alba personata	SGA1785	Altaiskii kray, Russia	52.38	85.67	M
M. alba personata	SGA1788	Altaiskii kray, Russia	52.37	85.54	M
M. alba personata	SGA1791	Altaiskii kray, Russia	52.37	85.54	M
M. alba personata	SGA1793	Altaiskii kray, Russia	52.35	85.34	M
M. alba personata	SGA1799	Altaiskii kray, Russia	52.28	85.44	M
M. alba personata	SGA1800	Altaiskii kray, Russia	52.22	85.47	M
M. alba personata	SGA1802	Altaiskii kray, Russia	52.22	85.47	M
M. alba personata	SGA1803	Altaiskii kray, Russia	52.22	85.47	M
M. alba personata	SGA1804	Altaiskii kray, Russia	52.22	85.47	M
M. alba personata	SGA1807	Altaiskii kray, Russia	52.22	85.47	M
M. alba personata	SGA1809	Altaiskii kray, Russia	52.22	85.47	M
M. alba personata	SGA1818	Altaiskii kray, Russia	52.45	85.42	M
M. alba personata	SGA1832	Altaiskii kray, Russia	52.56	85.34	M
M. alba personata	SGA1833	Altaiskii kray, Russia	52.56	85.34	M
M. alba personata	SGA1838	Altaiskii kray, Russia	52.17	85.86	M
M. alba personata	SGA1536	Uzbekistan	41.31	69.44	M
M. alba personata	SGA1539	Uzbekistan	41.31	69.44	M
M. alba personata	SGA1541	Uzbekistan	41.31	69.44	M
M. alba personata	SGA1542	Uzbekistan	41.31	69.44	M
M. alba personata	SGA1545	Uzbekistan	41.31	69.44	M
M. alba personata	SGA1546	Uzbekistan	41.31	69.44	M
M. alba personata	SGA1548	Uzbekistan	40.91	69.79	M
M. alba personata	SGA1550	Uzbekistan	40.95	69.83	M
M. alba personata	SGA1554	Uzbekistan	40.95	69.83	M
M. alba personata	SGA1561	Uzbekistan	40.95	69.83	M
F. albicollis	SAMEA3166414	Hynkov, Czech Republic	49.67	17.17	M
F. albicollis	SAMEA3166416	Hynkov, Czech Republic	49.67	17.17	M
F. albicollis	SAMEA3166417	Hynkov, Czech Republic	49.67	17.17	M
F. albicollis	SAMEA3166420	Hynkov, Czech Republic	49.67	17.17	M
F. albicollis	SAMEA3166421	Hynkov, Czech Republic	49.67	17.17	M
F. albicollis	SAMEA3166423	Hynkov, Czech Republic	49.67	17.17	M
F. albicollis	SAMEA3166427	Hynkov, Czech Republic	49.67	17.17	M
F. albicollis	SAMEA3166428	Hynkov, Czech Republic	49.67	17.17	M
F. albicollis	SAMEA3166429	Hynkov, Czech Republic	49.67	17.17	M
F. albicollis	SAMEA3166430	Hynkov, Czech Republic	49.67	17.17	M
F. hypoleuca	SAMEA3167980	Karlova, Studánka, Czech Republic	50.08	17.31	M
F. hypoleuca	SAMEA3167983	Karlova, Studánka, Czech Republic	50.08	17.31	M
F. hypoleuca	SAMEA3167986	Karlova, Studánka, Czech Republic	50.08	17.31	M
F. hypoleuca	SAMEA3167987	Karlova, Studánka, Czech Republic	50.08	17.31	M
F. hypoleuca	SAMEA3167990	Karlova, Studánka, Czech Republic	50.08	17.31	M
F. hypoleuca	SAMEA3167992	Karlova, Studánka, Czech Republic	50.08	17.31	M

F. hypoleuca	SAMEA3167994	Karlova, Studánka, Czech Republic	50.08	17.31	M
F. hypoleuca	SAMEA3167995	Karlova, Studánka, Czech Republic	50.08	17.31	M
F. hypoleuca	SAMEA3167997	Karlova, Studánka, Czech Republic	50.08	17.31	M
F. hypoleuca	SAMEA3167998	Karlova, Studánka, Czech Republic	50.08	17.31	M
F. parva	SAMEA3175208	Sweden	-	-	Unknown
F. hyperythra	SAMEA3175226	Indonesia	-	-	F

[a] MZUSP, Museu de Zoologia da Universidade de São Paulo; CUMV, Cornell University Museum of Vertebrates, KUNHM, University of Kansas Museum of Natural History; MCP, Coleção de Aves do Museu de Ciências e Tecnologia da Pontifícia Universidad Católica do Rio Grande do Sul; MACN, Museo Argentino de Ciencias Naturales "Bernardino Rivadavia," IOZ: Institute of Zoology Chinese Academy of Science, KU: Kansas University, UWBM: University of Washington Burke Museum, ZMUK: Zoologisches Museum der Christian-Albrechts-Universität, Kiel, Germany