2019-12-10

Genomic Analyses Reveal the Evolutionary Dynamics of Speciation in Two Tropical Avian Systems

Sarah Alison Cowles
University of Miami, scowles1@gmail.com

UNIVERSITY OF MIAMI

GENOMIC ANALYSES REVEAL THE EVOLUTIONARY DYNAMICS OF SPECIATION IN TWO TROPICAL AVIAN SYSTEMS

By

Sarah Alison Cowles

A DISSERTATION

Submitted to the Faculty
of the University of Miami
in partial fulfillment of the requirements for
the degree of Doctor of Philosophy

Coral Gables, Florida

December 2019

UNIVERSITY OF MIAMI

A dissertation submitted in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

GENOMIC ANALYSES REVEAL THE EVOLUTIONARY DYNAMICS OF SPECIATION IN TWO TROPICAL AVIAN SYSTEMS

Sarah Alison Cowles

Approved:

_____
J. Albert C. Uy, Ph.D.
Associate Professor of Biology

_____
William A. Searcy, Ph.D.
Professor of Biology

_____
Donald L. DeAngelis, Ph.D.
Research Associate Professor of Biology

_____
Kevin G. McCracken, Ph.D.
Associate Professor of Biology

_____
Darren E. Irwin, Ph.D.
Professor of Zoology
University of British-Columbia

_____
Guillermo Prado, Ph.D.
Dean of the Graduate School

COWLES, SARAH ALISON (Ph.D., Biology)
Genomic Analyses Reveal the Evolutionary Dynamics of Speciation in Two Tropical Avian Systems (December 2019)

Abstract of a dissertation at the University of Miami.

Dissertation supervised by Professor J. Albert C. Uy.
No. of pages in text. (140)

The process of speciation is long and complex, and takes place in multiple stages, with a variety of potential outcomes. Studying lineages at different stages in the speciation process can lend insight into the evolutionary factors that generate and maintain biodiversity. For my dissertation, I used two genomic datasets generated from field-collected DNA samples in two tropical avian systems to examine the early and late stages of the speciation process. I studied the initial stages of speciation in the Amazilia Hummingbird (*Amazilia amazilia*) subspecies complex. Found across Ecuador and Peru, this species is split into six phenotypically differentiated subspecies. I used a large genotyping-by-sequencing dataset and one mitochondrial marker to assess the phylogeny, population structuring, and gene flow between the subspecies. I found that the six subspecies split into three distinct clades corresponding with geography, with evidence of gene flow across some neighboring subspecies groups. In addition, using environmental data in concordance with the tree, I found that both expansion into new habitats and geographic isolation likely shaped the diversification of subspecies.

To study the final stages of speciation—in which two lineages come into secondary contact, I used two species of *Zosterops* White-eyes, *Z. kulambangrae* (the Solomons White-eye) and *Z. murphyi* (the Kolombangara White-eye). These two species are part of the Zosteropidae family of birds, which is known for rapidly speciating yet

still becoming geographically widespread (i.e., the paradox of the great speciator). Found on the island of Kolombangara in the Solomon Islands, *Z. kulambangrae* and *Z. murphyi* have an elevational contact zone at mid-elevations on the island. Using a large genotyping-by-sequencing dataset and one mitochondrial marker, I found an absence of hybridization and strong reproductive boundaries between the two species, even though the species have only been diverging for approximately two million years. I also explore potential mechanisms for reproductive isolation such as plumage and song. Putting my results into context with other studies using a literature review, I found that in comparison to other avian species pairs, these species have rapidly evolved complete reproductive isolation, which may help to explain the paradox of the great speciator.

Finally, I also use the *Zosterops* genomic dataset to predict current-day population sizes and structuring of both species on the island of Kolombangara, interpreting these population sizes in terms of historical logging and the needs for present-day conservation. Overall, in my dissertation I show the importance of using genomic data in concert with phenotypic and environmental datasets to study different stages of the speciation process. Studying speciation at various time points across different lineages is valuable for understanding how and why species are formed or go extinct, and the mechanisms responsible.

ACKNOWLEDGEMENTS

Of utmost importance, I would like to thank my dissertation advisor Al Uy for all of the help, guidance, financial support, and direction he has provided me over the last six years in developing as a scientist, and for helping me to focus on the "bigger picture" and overarching research goal. Thank you for taking me into your lab, introducing me to the beauty of the Solomon Islands, and for always being positive. I would also like to thank my committee member William Searcy for his level-headed thinking and helpful input in developing and writing up my dissertation. Thank you to my other committee members Don DeAngelis, who took so much time to help me develop potential research ideas, Kevin McCracken, for being a friendly face to talk to, and my external committee member Darren Irwin, who always provided wonderful feedback, encouragement, and insight from afar.

I would like to thank everyone who assisted with data collection and permitting in the field in both the Solomon Islands and Ecuador. In the Solomon Islands, I am particularly grateful to Jino Ghemu (my trustworthy and dedicated head field ranger on Kolombangara), Bryce Kanavari, Wesley Sparagon, and the many local field rangers from Kolombangara who aided in the often challenging and difficult fieldwork. I would like to thank the people of the Solomon Islands, particularly the warm and welcoming village of Iriri on Kolombangara, for their hospitality and care during my two field seasons on the island. I would also like to thank David Boseto and the staff of Ecological Solutions Solomon Islands for aiding in the logistics, permitting, and organization of the field trips. Many thanks to Ferguson Vaghi and the Kolombangara Island Biodiversity Conservation Association for allowing me to work on Kolombangara, and for helping to

organize the long field trips. Additionally, thank you to the Solomon Islands Ministry of Environment for permits to work in the beautiful Solomon Islands.

In Ecuador, I am particularly thankful for my field assistants David Anchundia, Javier Pinto, Laura Cremin, and Diego Ocampo, whose positivity and helpfulness made the field trips so much more enjoyable. Additionally, I would like to thank my collaborators Elisa Bonaccorso for aiding in getting the Ecuador hummingbird samples to the USA, and Chris Witt for giving us access to the Peruvian tissue samples and obtaining the Nazca Valley samples, and for their ideas and input on the study. I am also indebted to the staff of Bosque Protector Cerro Blanco Reserve and Fundacion Pro-Bosque, Las Tangaras Reserve, and Bellavista Reserve in Ecuador for aiding in the logistics of fieldwork. In addition, I would like to thank José Luis Ponce of Malacatos, Rodny Garrido of Mindo, Nicolas Ruhard of Vilcabamba, and the Ayampe and Cauchiche Communities for permission to use their lands for research in Ecuador. Additional thanks to the Ministerio del Ambiente de Ecuador for permits to conduct research across our field sites in Ecuador.

Thank you to my labmates over the years: Winter Beckles, José Hidalgo, Kristine Gandia, Emily Powell, Diego Ocampo, Jason Sardell, Doug Wiedemann, Dan Baldassarre, Liz Cooper, Erika Bueno, Krystle Young, and Floria Mora-Kepfer Uy for discussing research ideas, providing feedback, and for teaching me so many new things. A special thanks to Kristine, Emily, Winter, José, and Diego—thank you so much for your moral support, positivity, and friendship over the last six years. In addition, thank you to the former Searcy lab members: Karla Rivera-Caceres, Luis Vargas-Castro, Gavin

Leighton, and Adrienne DuBois for input on projects and stimulating discussion in our joint lab meetings.

A special thanks to all of the Department of Biology graduate students and faculty for their support and an enjoyable time in the department. I particularly would like to thank Alex Wilson and Athula Wikramanayake for all of the work they did in making the department a better place as graduate director and graduate chair, and for the financial support to attend the 2018 Workshop in Population and Speciation Genomics in the Czech Republic.

I thank my parents Rich and Jackie Cowles for allowing me to pursue my personal interests and goals and for their unwavering support for their crazy biologist daughter that would be incommunicado for weeks and sometimes months "in the jungle."

Finally**,** I thank my fiancé Felix Grewe for being my foundation and for keeping me focused on the end goal. I couldn't have done it without you.

The work in this dissertation was supported by many generous financial sources, including: a University Fellowship and Dean's Summer Research Fellowship from the College of Arts and Sciences at the University of Miami, the James W. McLamore Fellowship in Tropical Biology, a Tinker Foundation Field Research Grant from the University of Miami's Center for Latin American Studies, the University of Miami's Kushlan, Savage, and Evoy Biology Graduate Funds, the Lewis and Clark Fund from the American Philosophical Society, a Grant-in-Aid of Research from the Society for Integrative and Comparative Biology, a Van Tyne Research Award from the American Ornithological Society, and the Aresty Chair in Tropical Ecology at the University of Miami.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1

Introduction

"…from so simple a beginning endless forms most beautiful and most wonderful have been, and are being, evolved." -Charles Darwin, *On the Origin of Species* (1859)

The earth is amazingly diverse, and understanding how biodiversity is generated and maintained is central to the field of evolutionary biology (Darwin 1859, Mayr 1942, Gaston 2000, Coyne & Orr 2004). In particular, species are the fundamental units used to measure biodiversity. Although the definition of a species is debated, the most commonly used definition is that of a biological species. Dobzhansky (1935, 1937) and Mayr (1942) defined the biological species concept as "actually or potentially interbreeding populations that are reproductively isolated from other such groups." Speciation, or the process that gives rise to new species, can occur in several ways. With allopatric speciation, non-overlapping populations separated by physical barriers diverge over time such that if they come into secondary contact, they can no longer interbreed due to either pre-mating barriers such as divergent mating signals or post-mating barriers such as hybrid sterility or hybrid inviability. Allopatric speciation is pervasive and the most common form of speciation (Coyne & Orr 2004, Price 2008). Parapatric speciation, which occurs when two different populations only partially overlap, is theoretically possible and likely also occurs readily (Coyne & Orr 2004, Price 2008). Sympatric speciation, which is speciation in the absence of geographical barriers, is rare, and stringent criteria must be met to show that the species did not result from secondary contact after allopatry (Dieckmann & Doebeli 1999, Coyne & Price 2000, Savolainen et

al. 2006, Barluenga et al. 2006, Phillimore et al. 2008). Overall, the key feature of all modes of speciation is that some mechanism must reduce gene flow between populations, allowing for genetic isolation and divergence over time. In other words, speciation entails the evolution of reproductive isolation between two potentially interbreeding groups of organisms. However, the process of speciation is complex and takes place in stages, with a variety of potential outcomes.

Dobzhansky (1937) and Mayr (1942) proposed a multi-stage model of allopatric speciation. First, a phenotypically and genetically uniform species spreads over a large range and begins to diverge geographically. Second, isolating barriers arise, which result in population differentiation to the level of geographically variable species. Finally, due to range expansion or removal of barriers, two previously isolated forms come into secondary contact. Secondary contact can result in multiple outcomes, which can fall under two main categories: 1) weak reproductive isolation and the occurrence of hybridization (which can lead to a stable hybrid zone, fusion of the two species, hybrid speciation, or introgression of particular traits across the hybrid zone) or 2) strong reproductive isolation between the forms (which can result in no gene flow at all, or reinforcement that leads to no gene flow) (Dobzhansky 1937, Mayr 1942, Coyne & Orr 2004, Price 2008). Therefore, populations undergoing allopatric speciation can be classified into a particular stage of the speciation process (for example, initial splitting of lineages due to vicariance, or secondary contact resulting in a stable hybrid zone). Studying lineages at different stages in the speciation process can lend insight into the evolutionary factors that generate and maintain biodiversity.

In the last few decades, the advent of next-generation sequencing technologies has led to what is known as the "genomics revolution" in the field of biology (see Wolfe and Li 2003, Mardis 2008, Shendure and Ji 2008, Koboldt et al. 2013). Instead of relying on a few mitochondrial or microsatellite markers to explain evolutionary patterns, sequencing methods such as restriction site-associated DNA sequencing (RAD-seq) (see Andrews et al. 2011), targeted enrichment sequence capture (see Mamanova et al. 2013), and even full-genome sequencing (see Bentley 2006) have provided massive amounts of genomic data at an affordable cost. Particularly within the field of speciation, both genome-wide reduced representation single nucleotide polymorphism (SNP) datasets and whole-genome datasets have allowed for in depth studies of taxonomic relationships and secondary contact zones, particularly focusing on cases in which hybridization occurs (see Payseur and Rieseberg 2016 for a review). Many interesting phenomena like differential introgression of specific genomic regions or genes (e.g. Taylor et al. 2014, Toews et al. 2016), hybrid speciation (e.g. Barrera-Guzmán et al. 2017), and the detection of later generation backcrossed individuals (e.g. Oliviera et al. 2015, Scordato et al. 2017) can now be examined in more depth using these large scale genome-wide datasets.

In my dissertation, I used genomic datasets generated from two tropical avian systems to examine different stages of the speciation process. To study the initial diversification stage of speciation, I used the Amazilia Hummingbird (*Amazilia amazilia*). This neotropical species is found in Ecuador and Peru, and is split into six phenotypically variable subspecies across its geographic range. However, nothing was known about the evolutionary history, genetic differentiation, or level of gene flow across these subspecies prior to my study. To examine the final stages of speciation (secondary

contact), I used two species of *Zosterops* White-eye birds: *Z. kulambangrae* and *Z. murphyi*. These two species are found in the Paleotropics on the island of Kolombangara in the Solomon Islands, and have a secondary contact zone over an elevational gradient. I conducted the first detailed study of this contact zone using both genomic and phenotypic data. Below, I describe my two study systems in further detail, and then provide an outline of my following dissertation chapters.

Study Systems

The Amazilia Hummingbird (*Amazilia amazilia*) Subspecies Complex

The hummingbird family Trochilidae is extremely diverse, with approximately 340 species of hummingbirds classified into nine major clades (Bleiweiss et al. 1997, McGuire et al. 2014). Hummingbirds have radiated throughout most of North and South America, and the diversification of this clade has been aided by the uplift of the Andes, which created varied habitats along elevational gradients, as well as the formation of the Panamanian land bridge, which allowed for multiple cross-continental invasions by different clades (Graham et al. 2009, McGuire et al. 2014). Given the recent uplift of the Andes within the last 10 million years, and the formation of the Panamanian land bridge in the past 5 million years, much of the diversification in this family is quite recent (McGuire et al. 2014). This recent diversification makes hummingbirds an ideal system in which to study the initial stages of speciation, as many clades are still in the process of rapidly diverging. Particularly, the bee, emerald, and gem hummingbird clades have elevated net diversification rates, which are closely tied to their expansion into the Andes and invasions across the Panamanian land bridge (McGuire et al. 2014).

Within the emeralds, the *Amazilia* genus consists of over 29 medium-sized hummingbird species found from the USA down to Argentina that arose during the late Miocene and Pliocene (11.63-2.58 million years ago) (Ornelas et al. 2013, McGuire et al. 2014). Many species within the *Amazilia* clade are found over wide geographic ranges and split into multiple subspecies, with complex phylogeographic patterns, hybridization across species, and unclear species boundaries (e.g. Miller et al. 2011, Rodríguez-Gómez et al. 2013, Ornelas et al. 2014, Rodríguez-Gómez and Ornelas 2014, Rodríguez-Gómez and Ornelas 2015, Jiménez and Ornelas 2016, Rodríguez-Gómez and Ornelas 2018). The complexity of these taxonomic relationships and unclear species boundaries in this young clade warrants further study. However, no existing studies on the *Amazilia* genus have specifically taken advantage of a genomic-wide dataset to examine the underlying evolutionary patterns within a variable species complex.

Of particular interest is the Amazilia Hummingbird (*Amazilia amazilia*). This species is spread over a wide geographic range, and is found from the Colombian-Ecuadorian border down to the Nazca Valley in Peru on the western side of the Andes. Split into six distinct subspecies that vary phenotypically in plumage characteristics, range size, and in the degree of geographic isolation (Weller 2000, Krabbe and Ridgely 2010), this system lends itself well to a study examining the initial stages (diversification) of the speciation process in a geographically variable species using a genome-wide dataset.

The Kolombangara White-eye (*Zosterops murphyi*) and Solomons White-eye (*Z. kulambangrae*)

The Zosteropidae family of birds consists of approximately 115 species found in 14 genera, and 75% of the species fall within the *Zosterops* genus, which is the most species-rich genus of birds (van Balen 2001). Much of the diversity within this clade has arisen within the last 2 million years (Moyle et al. 2009). Species in this family are found across Africa, Asia, through the Indo-Pacific (see Fig. 1.1). Classified as a "great speciator" (Diamond et al. 1976), this lineage has simultaneously rapidly speciated, yet spread over a wide geographic range. This idea is counterintuitive, as speciation necessarily requires the evolution of reproductive isolation (and thus a lack of gene flow), but dispersal across a wide geographic range necessarily implies high levels of migration and the potential for subsequent gene flow.

In particular, this family of birds has one of the highest per-lineage diversification rates known for vertebrates (Moyle et al. 2009, Jetz 2012), with a per-lineage diversification rate estimated at 1.95-2.63 new species per million years (Moyle et al. 2009). Compared to other rapidly speciating lineages with similar diversification rates such as Darwin's finches, Hawaiian honeycreepers, and African cichlids that are found in isolated regions (with a variety of open niches to fill), Zosteropidae is found over a much wider geographic range, and documentation exists of its increased dispersal ability through natural range expansions (North 1904; Falla et al. 1966; Mees 1969; Clegg et al. 2002, Melo et al. 2011, Cox et al. 2014). Therefore, Zosteropidae lends itself to in-depth studies of speciation and secondary contact. Several examples exist where two sympatric or parapatric *Zosterops* lineages are thought to have come into secondary contact (e.g

Warren et al. 2006; Clegg and Phillimore 2010; Melo et al. 2011; Wickramasinghe 2017, O'Connell et al. 2019); however, only two *Zosterops* contact zones have been studied in depth using genetic markers. One zone is on the island of Reunion (Gill 1970, 1973, Milá et al. 2010); the other is in the Cape Region of South Africa (Oatley et al. 2012, 2017). In both cases, microsatellite markers and plumage characteristics have shown that some hybridization has occurred in these zones, but these zones are quite young in age (<1 million years old). Given the Zosteropidae lineage's wide geographic range and increased dispersal ability, secondary contact zones in *Zosterops* remain highly understudied within a genomic context—particularly within contact zones greater than 1 million years old, or through using a large genomic-wide dataset. In addition, little is known about the specific reproductive barriers that contribute to reproductive isolation between species.

On the island of Kolombangara in the New Georgia Province of the Solomon Islands, two overlapping *Zosterops* species exist. One species, *Z. kulambangrae*, is found on lower elevations below 1000 meters, whereas the other species, *Z. murphyi*, is endemic to the island at elevations greater than 600 meters. These two species have a naturally occurring secondary contact zone at mid-elevations, in which they are found together in mixed-species flocks. These two species are not sister to one-another and have likely evolved within the last two million years from independent invasions (Mayr and Diamond 2001, Moyle et al. 2009). Therefore, the two species on Kolombangara provide an ideal system in which to study the genomic consequences (presence or absence of gene flow) of an older secondary contact zone within Zosteropidae using a genomic-wide dataset, and to examine potential reproductive barriers.

Dissertation Outline

In the following chapters of my dissertation, I present three studies in which I used genomic-wide data combined with either phenotypic or environmental data to explore the evolutionary dynamics of speciation and conservation in two tropical avian systems. In Chapter 2, I examine the initial stages of the speciation process in the *Amazilia amazilia* hummingbird subspecies complex. I collected DNA samples from multiple field sites throughout the species range in Ecuador, and obtained museum genetic tissue samples from across the species range in Peru. Using one mitochondrial DNA marker and a large genotype-by-sequencing (GBS) single nucleotide polymorphism (SNP) dataset, I reconstructed the phylogeny and evolutionary history of the subspecies, examined population structuring, and tested for gene flow among neighboring subspecies. In addition, I used environmental data in concert with the phylogeny to assess whether expansion into new habitats or geographic isolation shaped the present-day distribution of subspecies.

Chapter 3 examines the consequences of secondary contact between the two species of *Zosterops* White-eyes on the island of Kolombangara in the Solomon Islands. I captured birds from both species along two elevational transects from different sides of the island. Using a mitochondrial marker and a large GBS SNP dataset, I assessed whether gene flow was occurring between the two species. In addition, I examined song and plumage as potential pre-mating barriers to gene flow between these two species. Finally, I placed my results into a broader context by conducting a literature review to examine the time it takes to achieve reproductive isolation (i.e., a reduction of gene flow)

in two congeneric avian species or subspecies (with a parapatric distribution or known hybrid zone) in secondary contact.

In Chapter 4, I used the genomic dataset generated in Chapter 3 to estimate the current effective population sizes of the two species of *Zosterops* White-eye birds on Kolombangara using two different programs designed to work with genomic SNP data. I also examined fine-scale population structuring on the island and assessed relatedness among individuals. Given that *Z. murphyi* is endemic to the island whereas *Z. kulambangrae* is found over a wider range across the New Georgia Islands, I interpret the results in terms of conservation, and assess whether large-scale logging on Kolombangara Island has affected population numbers of the two species.

Finally, in Chapter 5 I summarize the overall findings of my dissertation research and place the results into a broader context within the field of evolutionary biology. Across all chapters, I applied genomic approaches in concert with fieldwork, lab work, phenotypic analyses, and use of environmental data to examine different stages of the speciation process and conservation in birds.

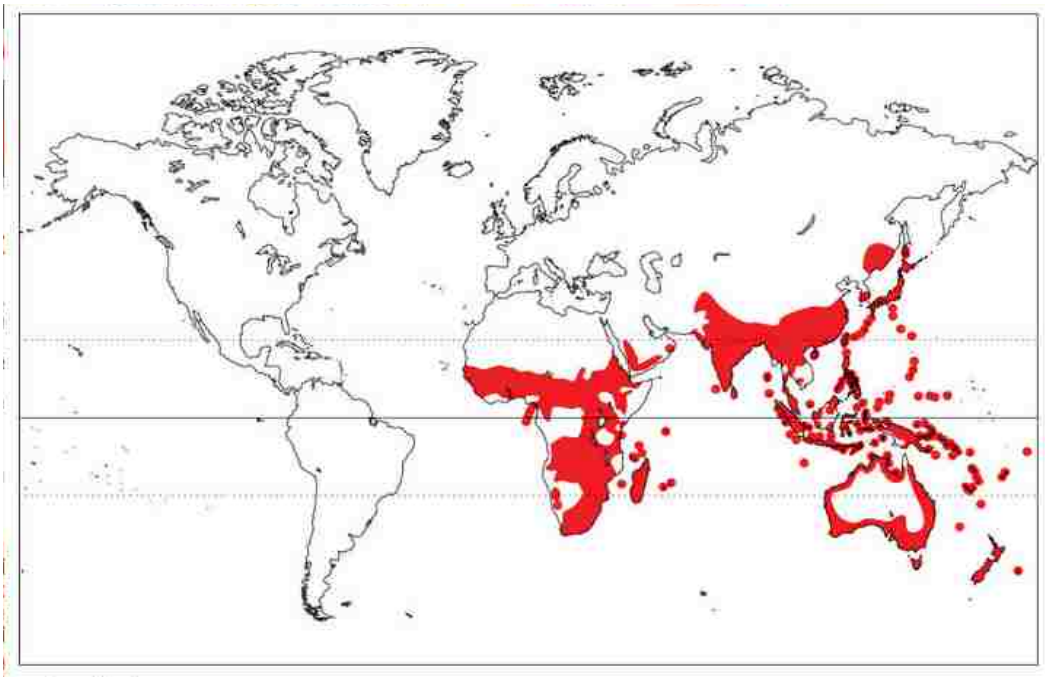

Figure 1.1: Species in the Zosteropidae family of birds cover a wide geographic range (denoted in red) and are found in Africa, Asia, Australia, and the Indo-Pacific. Figure taken from the Internet Bird Collection.

Chapter 2

Range expansion and geographic isolation mediate the diversification of the Amazilia Hummingbird (*Amazilia amazilia*) subspecies complex

Summary

Disentangling the factors underlying the diversification of a species with a wide geographical range is essential to understanding the initial stages and drivers of the speciation process. The Amazilia Hummingbird, *Amazilia amazilia*, is found along the Pacific coast and highlands from the Ecuadorian-Colombian border down to the Nazca Valley of Peru, and is currently classified into six phenotypically differentiated subspecies. Our aim was to resolve the evolutionary history of the six subspecies, to assess population structure, and to determine if introgression is taking place using a genome-by-sequencing dataset of 86 individuals from across the species range. The consensus phylogenetic tree separated the six subspecies into three distinct clades, corresponding with the Ecuador lowlands (*A. a. dumerilii*), the Ecuador highlands (*A. a. alticola* and *A. a. azuay*), and the Peruvian coast (*A. a. leucophoea*, *A. a. amazilia*, and *A. a. caeruleigularis*). We also found evidence of gene flow between the subspecies *A. a. dumerilii*, *A. a. alticola*, and *A. a. leucophoea*, with strong genetic isolation of the subspecies *A. a. azuay* in the isolated Yunguilla Valley of Ecuador. Finally, environmental data from the ranges of each subspecies is concordant with the three distinct clades. Overall, our results suggest that both expansion into new habitats and geographic isolation shaped the present-day phylogeny and range of the *A. amazilia* subspecies.

Background

Many factors can shape the evolutionary trajectory of a species. Environmental features are particularly important in divergence, as they can become barriers to or facilitate gene flow, lead to ecological specialization, and affect community structure (e.g. Dobzhansky 1937, Coyne and Orr 2004, Rundle and Nosil 2005, Price 2008, Graham et al. 2009, Nosil 2012). Mountain ranges, such as the Andes, are among the most important geographic barriers and are known to be incredibly biodiverse (Körner and Spehn 2002). The Andes are the longest above-sea mountain chain in the world, spanning almost the entirety of the Pacific side of South America from Venezuela to Argentina. Pleistocene glacial cycles greatly influenced biodiversity in the Andes, as contraction of montane habitats along elevational gradients during warming periods allowed for the isolation required for speciation, with subsequent expansion of montane habitats during cooling allowing dispersal into new areas. Therefore, many rapidly diversifying lineages have been associated with the geology of the Andes, in both plants (e.g. Hughes and Eastwood 2006, Pérez-Escobar et al. 2017) and animals (e.g. Elias 2009, Weir and Price 2011, Hutter et al. 2013, Beckman and Witt 2015).

Within the Americas, the hummingbird family (Trochilidae) has undergone a massive diversification, radiating into approximately 340 species of hummingbirds that are classified into nine major clades (Bleweiss et al. 1997, McGuire et al. 2014). Much of this family's diversity has been shaped by landscape features such as the uplift of the Andes or formation of the Panamanian land bridge. These geographical features allowed for isolation and provided a variety of niche habitats across elevational gradients, which aided in the diversification of this lineage (Graham et al. 2009, McGuire et al. 2014).

Hummingbirds invaded South America by approximately 22 million years ago, and rapidly diversified into the Andes. Approximately 40% of all hummingbird species are now found within the Andes (McGuire et al. 2014).

Net diversification rates vary among clades in hummingbirds, making them an ideal system to study the underlying processes that drive speciation. Several clades, including bee hummingbirds, mountain gems, and emeralds, show high net diversification rates, which are likely tied to expansion into new ranges during the uplift of the Andes and the formation of the Panamanian land bridge (McGuire et al. 2014). The clade of emerald hummingbirds contains 108 species and has evolved within the last 10-15 million years, with a net diversification rate estimated at 0.33 species/lineage/million years (McGuire et al. 2014). Within the emerald clade is the *Amazilia* genus, which contains over 29 species of medium-sized hummingbirds that are distributed from the Southern USA to Argentina. The *Amazilia* genus likely originated in the Isthmus of Tehuantepec and diversified in two major clades, with one clade spreading east of the isthmus and one clade spreading to South America, with most diversification occurring between the late Miocene and Pliocene (11.63-2.58 million years ago) (Ornelas et al. 2014, McGuire et al. 2014*)*. Many studies have evaluated the complex phylogeographic patterns, hybridization patterns, and drivers affecting species and subspecies diversification in this diverse genus across Central and South America (e.g. Miller et al. 2011, Ornelas et al. 2013, Rodríguez-Gómez et al. 2013, Rodríguez-Gómez and Ornelas 2014, Rodríguez-Gómez and Ornelas 2015, Jiménez and Ornelas 2016, Rodríguez-Gómez and Ornelas 2018), which are consistent with a young, rapidly diversifying clade. However, no studies in the *Amazilia* genus have yet taken advantage of a genotype-by-

sequencing dataset to examine the diversification of a variable subspecies complex in depth within this clade.

The Amazilia Hummingbird (*Amazilia amazilia*) is a species of medium-sized hummingbird (9-10 cm, 4-7g) that is found along the western coast of Ecuador from close to the Ecuadorian-Colombian border down to the Nazca Valley in Peru. These hummingbirds inhabit arid and semi-arid lowland scrub/dry forest environments along the Pacific coast and can also range up into the subtropical forest on the Andean slopes to elevations of up to 2800 meters. In addition, they can be found within the gardens of towns and cities, such as Lima and Guayaquil (Calvino-Cancela 2006, Weller et al. 2019). *A. amazilia* feeds on nectar from flowers of medium corolla length such as *Salvia splendens*, *Justicia brandegeana*, *Erythrina*, *Psittacanthus*, and *Leonotis nepetifolia*, as well as on small insects that are caught aerially (Calvino-Cancela 2006, Weller et al. 2019, S. Cowles pers obs). This species is also territorial against conspecifics, other species of hummingbirds, and other nectar feeders such as banaquits (Calvino-Cancela 2006, S. Cowles pers obs), and may show small-range altitudinal migrations following food sources across seasons (Weller et al. 2019, S. Cowles pers obs). The current classification recognizes six distinct subspecies—three in Ecuador: *A. amazilia alticola*, *A. amazilia azuay*, and *A. amazilia dumerilli*, and three in Peru: *A. amazilia leucophoea*, *A. amazilia amazilia*, and *A. amazilia caeruleigularis* (Fig. 2.1; from here onward subspecies will be written only using the subspecies designation, i.e. *alticola*, *azuay*, *dumerilii*, *leucophoea*, *amazilia*, and *caeruleigularis*). These subspecies differ remarkably in several phenotypic characters, such as the presence or absence of white throat patches, rufous belly coloration, tail coloration, and gorget coloration (Fig. 2.1;

Weller 2000, Krabbe & Ridgely 2010), making them ideal to study the factors that underlie divergence in the early stages of the speciation process.

Previous research has suggested that the northern (*dumerilii* and *leucophoea*) and southern (*amazilia* and *caeruleigularis*) subspecies of *A. amazilia* form separate clades due to phenotypic similarity, and a south to north dispersal aided subspecies diversification due to phenotypic similarity with closely-related congeners (Weller 2000). However, the evolutionary relationships and history of the higher altitude subspecies in Ecuador (*alticola* and *azuay*) in relation to these possible northern and southern clades remain unknown, and there is debate on whether the subspecies *alticola*, known as the Loja Hummingbird, should be elevated to species status (Weller 2000, Krabbe and Ridgely 2010). In addition, little is known about the level of hybridization or intergradation between the six subspecies of *A. amazilia* across Ecuador and Peru.

The aims of our study were to first examine the phylogenetic relationships of all six subspecies of *A. amazilia*, and then to assess the level of gene flow and population structuring across the six subspecies throughout their range using a variety of genomic techniques. In our analyses, we used genomic DNA isolated from 86 field-collected blood samples and tissue samples from museum specimens to create a genomic-wide dataset of SNPs. In addition, using climate data from each subspecies range, we explore geographical boundaries and ecological barriers that potentially led to the subspecies diversification within the *A. amazilia* group.

Methods

Field Methods (Ecuador Samples)

We conducted fieldwork from May through July of 2014 to capture *A. amazilia* in Ecuador. We captured birds from six different field sites around the range of the three Ecuadorian subspecies (*alticola*, *azuay*, and *dumerilii*; see Table 2.1 for details). To catch birds, we used both six- and 12-meter mist nets in flyways and close to nectar sources, and occasionally used red nectar feeders filled with sugar water as bait. Once captured, we took blood samples from the medial metatarsal vein of each hummingbird. We also clipped an outer right tail retrix to identify the individual in case of recapture. Each bird was given sugar-water ad libitum and subsequently released. Over our field season, we captured a total of 55 *A. amazilia* across our six locations in Ecuador (Table 2.1). All field methods were approved beforehand by the University of Miami's Institutional Animal Care and Use Committee.

In addition to capturing *A. amazilia*, we also captured individuals of two congeneric species at two field sites: the Rufous-tailed Hummingbird (*A. tzacatl*, 9 samples) and the Andean Emerald (*A. franciae*, 2 samples) (see Table 2.1 for details). Given that the phylogeny of the entire Mesoamerican *Amazilia* genus is resolved (Ornelas et al. 2013, McGuire et al. 2014), we used samples from these two species as outgroups in our subsequent phylogenetic analyses. We extracted DNA from all *A. amazilia* and congeneric samples using Qiagen's DNeasy Blood and Tissue Extraction Kits (Qiagen, Hilden, Germany) using the manufacturer's protocol for nucleated blood.

Peruvian Tissue Samples

We obtained 34 tissue samples of *A. amazilia* from the Museum of Southwestern Biology at the University of New Mexico (see Table 2.1) that spanned the species range in Peru (West of the Andes from Tumbes/Piura down to the Nazca Valley; including subspecies *leucophoea*, *amazilia*, and *caeruleigularis*). Tissue samples were collected between 2001-2017 and consisted of a mix of heart, liver, and muscle samples that had been previously frozen and/or stored in RNA later. We extracted DNA using Qiagen's DNeasy Blood and Tissue Extraction Kits (Qiagen, Hilden, Germany) using the manufacturer's protocol for tissue samples, with the suggested addition of 4ul of RNase A per sample after tissue digestion to reduce the possibility of any RNA contamination.

Mitochondrial Sequencing and Analysis

Our aim was to sequence the mitochondrial gene NADH dehydrogenase subunit 2 (*ND2*). We first amplified the H1064 and L5215 primers using a standard PCR protocol (Sorenson et al. 1999) with an annealing temperature of 54°C. We cleaned the PCR-amplified DNA using the manufacturer's ExoSAP-IT protocol (USB corporation, Cleveland, OH, USA), and prepared the DNA for sequencing using the BigDyeTerminator Version 3.1 Cycle Sequencing Kit (Applied Biosystems, Waltham, MA, USA). DNA was purified before sequencing using Sephadex (Sigma-Aldrich, St. Louis, MO, USA) and was loaded onto 96-well plates. Plates were sequenced using Sanger Sequencing in the Institute of Biotechnology at Cornell University (Ithaca, NY, USA). Mitochondrial sequences were then trimmed, inspected, and aligned using Sequencher version 5.4.6 (Gene Codes, Ann Arbor, MI, USA). Finally, we used the

software program PopART version 1.7 (Leigh and Bryant 2015) to create a median joining haplotype network for all samples.

Genomic Sequencing

Our aim was to obtain thousands of single nucleotide polymorphisms (SNPs) across the genome using the reduced representation sequencing method of genotyping-by-sequencing (GBS) (Elshire et al. 2011). We chose 86 of our 89 *A. amazilia* samples based on extracted DNA concentrations (all but MSB 34693, MSB 32901, and MSB 43305), 7 *A. tzacatl* (3 from Mindo and 4 from Ayampe*)*, and the 2 *A. franciae* samples to fill a 96-well plate (with one blank well as a control). We first checked extracted genomic DNA quality by running 100ng of each sample on a 1% agarose gel to confirm the presence of intact, non-fragmented genomic DNA. We then digested 10% of the samples using the manufacturer's EcoRI digestion protocol (Promega, Fitchburg, WI, USA) to confirm the presence of enzymatic activity. We loaded samples onto a 96-well plate and sent the plate to the University of Wisconsin-Madison's Bioinformatics Resource Center for GBS using the restriction enzyme ApeKI. SNPs were called by the University of Wisconsin-Madison's Bioinformatics Resource Center using the established UNEAK-TASSEL pipeline and parameters (Glaubitz et al. 2014). This UNEAK-TASSEL pipeline calls SNPs with a minimum minor allele frequency of 0.05, a minimum count of 5, a mismatch rate of below 0.03, and a minimum call rate of 0.1. This preliminary filtering resulted in a dataset of 1,032,375 SNPs. We additionally filtered the dataset by choosing only SNPs that had a minor allele frequency of at least 0.05 and that were found in at least 60 of the 96 sequenced individuals. This reduced dataset ultimately

included 34,896 SNPs, and we created a vcf file of these SNPs in TASSEL version 1.5 (Bradbury et al. 2007) to use in our downstream analyses.

Genomic Analyses

To obtain estimates of population structuring and gene flow across the six subspecies of *A. amazilia*, we used several genomics programs and methods designed to work with genomic SNP data. To reconstruct a phylogeny, we first converted our vcf file of 34,896 SNPs into an interleaved phylip file in TASSEL version 1.5 (Bradbury et al. 2007). We then used the program RAxML (Randomized Axelerated Maximum Likelihood) version 8.2.8 (Stamatakis 2014) to build a maximum likelihood tree of the 86 *A. amazilia* samples and the two congener species (7 *A. tzacatl* and 2 *A. franciae* samples). We used the general time reversal model with gamma correction (GTRGAMMA) and the rapid bootstrap option (Stamatakis 2008) to replicate 100 trees. The resulting phylogenetic tree was then mid-point rooted and drawn to scale in MEGA version 7.0 (Kumar 2016).

Next, we conducted a principal component analysis (PCA) of our genomic data in TASSEL version 1.5 (Bradbury et al. 2007) after removing the congener taxa from the vcf file. We first imputed missing values for the PCA using the 'LD-kNNi' method in TASSEL (see Money et al. 2015) with the following parameters: high LD Sites = 30 (default), number of nearest neighbors = 3, and the max distance between sites for LD = 100,000. We then ran the PCA and plotted the eigenvalues for PC1 vs PC2 for each individual in R v3.3.2 (R Development Core Team 2016).

We also used the Bayesian program fastStructure (Raj et al. 2014) to examine population structuring and grouping of individuals. We first rearranged our vcf file with 34,896 SNPs and the 86 *A. amazilia* individuals (after removing the congener taxa) to cluster individuals at the subspecies level according to our RAxML tree (i.e., *alticola*, *azuay*, *dumerilii*, *leucophoea, amazilia*, and *caeruleigularis)* and then converted the vcf file into plink bed format using the program PLINK version 1.07 (Purcell et al. 2007). Next, we input the bed, bim, and fam files into fastStructure, and ran structure analyses using the available 'structure.py' script. In separate model runs, we changed the number of assigned populations (K) from 2 to 6. We used a convergence criterion of $10e^{-6}$, and the simple prior (flat-beta prior). We then used the 'chooseK.py' script available in fastStructure, which uses a marginal likelihood estimate on multiple model runs with different K values to choose the best-fit model to our dataset. To visualize our output meanQ files, we used the program Pophelper (Francis 2016; available from www.pophelper.com) to illustrate population clustering under different values of K.

To further examine fine-scale population structuring and clustering, we used the program fineRADstructure (Malinsky et al. 2018) to visualize estimated coancestry levels across individuals. fineRADstructure consists of several scripts to calculate coancestry between pairwise individuals (RADpainter), and subsequent clustering and population structuring between individuals (fineSTRUCTURE) using genomic SNP data. We first converted our subspecies-organized vcf file of 34,896 SNPs to the haplotype file input format for the program using the script 'hapsFromVCF' in RADpainter. Next, we calculated the coancestry matrix for all individuals using RADpainter, and then used the fineSTRUCTURE Markov chain Monte Carlo (MCMC) clustering algorithm with the

input arguments of x 100,000, -z 100,000, and -y 1,000. This MCMC algorithm repeatedly explores merging and splitting populations as well as moving individuals until a configuration is accepted that has a probability derived from the ratio of the likelihood from the previous configuration (see Lawson et al. 2012). This probabilistic process is repeated for each pair of individuals twice (each as the donor and as the recipient in the pair); hence, diagonal halves of the matrix may not be symmetrical. We then used the tree-building algorithm in fineSTRUCTURE with the input arguments of -m T -x 10000 to create a simple tree of the individuals. Finally, we visualized the program output with the fineSTRUCTURE GUI (available from https://people.maths.bris.ac.uk/~madjl/finestructure/ finestructure.html). Because the mcmc clustering and tree-building algorithm did not provide a useful tree (i.e., did not cluster individuals of the same subspecies or geographical location together in a logical manner or in accordance with our RAxML tree), we disregarded the output tree and instead visualized the RADpainter coancestry matrix (chunks.out file) with individuals of each subspecies clustered together.

To test for potential gene flow across subspecies, we used the program Dsuite (Malinsky 2019, available from github: https://github.com/millanek/Dsuite) to calculate the Patterson's D-statistic for subspecies trios, which is the test statistic for the ABBA-BABA test (Green 2010, Durand et al. 2011). The ABBA-BABA test uses the idea that phylogenetic trios can be used to test for the presence of introgression in non-sister taxa. Given a phylogeny with an outgroup O, and three taxa P1, P2, and P3 with a tree topology of (O, (P3, (P2, P1))), and alleles A and B, we would expect an equal number of (A, (B, (B, A)))s and (B, (A, (B, A)))s across the tree without introgression due to

random lineage sorting. However, with introgression between lineages P3 and P2, we would expect an abundance of the ABBA over a BABA pattern due to gene flow between P3 and P2. The D-statistic takes the ratio of the difference of ABBAs and BABAs over the total number of sites: D = (# of ABBAs - # of BABAs)/ (# of ABBAs + # of BABAs). The program then tests whether the D-statistic for each possible trio is different from 0 using a standard block jackknifing procedure (see Malinsky 2019), calculates a Z-score, and reports the associated p-values. Dsuite takes a vcf file as input, with the option to provide an input tree. We used a vcf file containing the 34,896 SNPs of all the *A. amazilia* samples and the two *A. franciae* samples specified as an outgroup. We input our resulting RAxML tree file (see above for methods, results for RAxML tree) into the program. We report results with a jackknifing parameter of 1000 (i.e., -j 1000; this parameter is supposed to be larger than the extent of linkage disequilibrium (Durand et al. 2011)). We tested a range of other values from 100-10000, but they did not qualitatively change the results). Based on the input tree, Dsuite assessed 20 different possible subspecies trios for introgression; however, of the 20 trios, only 13 trios fall under the five possible introgression scenarios that match the geography of the subspecies (e.g. testing introgression between *azuay* and *caeruleigularis* would not make sense geographically) and where the two taxa are sister or are already the most closely related to one another in the tree (i.e., we cannot assess *azuay-alticola* or *amazilia-caeruleigularis* for introgression since they are already the most closely-related taxa to one another in the tree) (Fig. 2.7). We used a Bonferroni correction on the p-values as suggested in Malinsky (2019) to account for multiple hypothesis testing.

As a final test of potential gene flow between subspecies, we used the program TreeMix v.1.13 (Pickrell and Pritchard 2012), which explores the possibility of migration events (and therefore shared ancestry) between non-sister subspecies within the tree. First, we converted our vcf file of the 86 *A. amazilia* samples into the TreeMix format using the script 'vcf2treemix.py' in the RAD_Tools github package (Baxter et al. 2011; available from https://github.com/CoBiG2/RAD_Tools/blob/master/vcf2treemix.py). We then ran the TreeMix program for the six subspecies, with 1000 SNP blocks (-k 1000) to account for any linkage disequilibrium, and for five migration events (-m parameter). We also included the basic tree topology with no branch lengths (i.e., from our RAxML tree previously) in our program runs (-tf parameter). We then plotted the results in R using the 'plotting_funcs.R' script provided in the src folder of the TreeMix program, and reported migration edge weights for each event. The migration edge weight represents the percentage of ancestry in the second lineage that is derived from the migration event (Pickrell and Pritchard 2012). Finally, we ran all possible 3-population and 4-population tests as recommended by the TreeMix authors and implemented within the TreeMix program (see Keinan et al. 2007, Reich et al. 2009 for statistical details) to help interpret our results. In our tests, we specified a block size of 500 SNPs (-k 500).

Habitat Differentiation

An additional aim of our project was to compare climatic variables (i.e., temperature and precipitation) in the range of each subspecies to determine whether geographic barriers (i.e., similar climatic environments across subspecies with divergence caused by geographic isolation) or ecological expansion into new habitats (i.e., different

climatic environments across subspecies) underlie the diversification of the *A. amazilia* subspecies. We used two different methods for this comparison using data extracted from the Worldclim2 database (Fick and Hijmans 2017) in R using the 'raster' package (Hijmans and van Etten 2014), using a spatial resolution of 2.5 minutes. First, we used two variables: annual mean temperature (variable 1) and annual precipitation (variable 12) from the Worldclim2 database for each of our GPS locations (see Table 2.1). To compare these two variables across subspecies, we conducted Kruskal-Wallis tests (and post-hoc pairwise Dunn tests, using the package 'dunn.test' (Dinno 2017)) in R version 3.3.2 (R Development Core Team 2016) to test if the mean temperature and annual precipitation in the habitats of each subspecies were significantly different across subspecies. For our second method, we conducted a PCA on all 19 of the Worldclim2 variables in R, and assessed whether the principal components were significantly different across subspecies using Kruskal-Wallis tests (and post-hoc pairwise Dunn tests when necessary).

Results

mtDNA

We were able to sequence and align 1002 base pairs of the mitochondrial gene *ND2* for a total of 75 *A. amazilia* individuals (17 *alticola*, 11 *azuay*, 18 *dumerilii,* 20 *leucophoea*, 5 *amazilia*, and 4 *caeruleigularis*). The median-spanning haplotype network shows that the mtDNA haplotypes of 5 of the 6 subspecies (all but *caeruleigularis*) are intermixed and not well-resolved (Fig. 2.2), and individuals from these 5 subspecies share a common mtDNA haplotype (23 individuals total, Fig. 2.2). The only distinct

mtDNA haplotype is that of *caeruleigularis*, which is separated from the common haplotype by 10 sequences changes, which is a sequence divergence of 1%.

We were also able to sequence and align 6 *ND2* sequences from the congener *A. tzacatl* to the 1002 base pair *A. amazilia ND2* sequences. There were a total of 101 sequence differences between the common haplotypes of *A. tzacatl* and *A. amazilia* out of 1002 total base pairs, which is a sequence divergence of approximately 10%. Similarly, we were able to sequence 890 base pairs from our 2 samples from *A. franciae*. We aligned these sequences to the *A. amazilia* sequences, and found a total of 58 differences across the 890 base pairs of *ND2,* which is a sequence divergence of approximately 6.51%.

To obtain estimates of divergence times between *A. amazilia* and the two congener species *A. franciae* and *A. tzacatl*, we used a standard mtDNA clock of 2% divergence per million years that is commonly used to estimate divergence time in birds (e.g. Arbogast et al. 2006, Price 2008, Cowles and Uy 2019). Using this clock rate, we estimate that *A. franciae* and *A. amazilia* last shared a common ancestor about 3.25 million years ago, whereas *A. tzacatl* and *A. amazilia* last shared a common ancestor about 5 million years ago. To better compare our results with Ornelas et al. (2013), which is the most complete *Amazila* genus phylogeny to date, we also used their unconventional (slow) clock rate of 0.0068 base substitutions per million years for the mtDNA gene *ND2*, which would estimate that *A. franciae* and *A. amazilia* last shared a common ancestor about 9.6 million years ago, whereas *A. tzacatl* and *A. amazilia* last shared a common ancestor about 15 million years ago. Finally, given that *caeruleigularis* has a

sequence divergence of 1%, we can estimate *caeruleigularis* has been isolated from the other subspecies for approximately 500,000 years (using a 2% divergence rate).

Phylogeny

Our RAxML maximum-likelihood phylogeny was based on a matrix of 34,896 SNPs from 86 *A. amazilia* and 9 congener DNA samples (7 *A. tzacatl* and 2 *A. franciae*), and had 31.49% gaps (percentage of missing data). The phylogeny shows clear separation of the six subspecies with high bootstrap support (Fig. 2.3, only bootstrap values $\geq$ 70 are shown). The samples fall into three distinct clades: a highlands Ecuador clade (subspecies *azuay* and *alticola*), a Peruvian clade (subspecies *leucophoea*, *amazilia*, and *caeruleigularis*), and a lowland Ecuador group (subspecies *dumerilii*). Based on branch lengths, *azuay* is the most genetically-distant subspecies, followed by *caeruleigularis* and *amazilia*. *Azuay* appears to have arisen from the secondary isolation of individuals from the *alticola* group. Similarly, *amazilia* and *caeruleigularis* (sister to one another) appear to have arisen from the *leucophoea* lineage.

Population Structuring

The PCA in the R package adegenet supports that *azuay* is the most genetically distinct subspecies, as the first principal component explained 7.21% of the genomic differentiation and cleared separated *azuay* from the other 5 subspecies (Fig. 2.4). The second principal component explained 4.25% of the genomic variation, and separated *caeruleigularis* and *amazilia* from *alticola*, *dumerilii*, and *leucophoea* (Fig. 2.4). However, individuals of *alticola*, *dumerilii*, and *leucophoea* do show some overlap in

both PC1 and PC2 (Fig. 2.4), suggesting they are genetically similar and may experience contemporary gene flow or have had recent gene flow.

Using the program fastStructure, we found that individuals of *A. amazilia* were clustered into distinct clades based on subspecies groupings in all model runs. With the number of specified populations at K = 2, *azuay* individuals were clustered as a separate genetic group from all other *A. amazilia* subspecies (Fig. 2.5). With K = 3, both *azuay* individuals and a single cluster of the *amazilia*/*caeruleigularis* group were identified as two distinct genetic groups, and then *alticola*, *dumerilii*, and *leucophoea* were all defined as an additional genetic group. In values of K = 4 to 6, fastStructure found four distinct genetic clusters: *azuay*, *amazilia*/*caeruleigularis*, and then variations of *dumerilii*, *alticola*, and *leucophoea* groupings to make two additional clusters (i.e., *alticola*/*dumerilii* and *leucophoea*, or *alticola*/*leucophoea* and *dumerilii*) (Fig. 2.5). We found that K= 4 was the best-fit model for explaining the structure in the data using the fastStructure script 'choose.py' for runs $K \geq 4$.

We used the fineRADstructure program to calculate pairwise coancestry values between all *A. amazilia* individuals. We found that the program showed higher coancestry values and structuring for both *azuay* and *caeruleigularis*, but otherwise detected little to no population structuring across the other four subspecies (*dumerilii*, *alticola*, *leucophoea,* and *amazilia*) (Fig. 2.6).

Introgression in non-sister subspecies

We used the program Dsuite to calculate D-statistics for possible introgression between non-sister subspecies trios. Over our five possible scenarios of introgression

between non-sister subspecies (based on the phylogenetic tree and geography, see Fig. 2.7), we found support for introgression in all cases (Table 2.2): *alticola-dumerilii* (3 of 4 trios), *azuay-dumerilii* (2 of 2 trios), *dumerilii-leucophoea* (2 of 3 trios), *alticola-leucophoea* (2 of 3 trios), and *amazilia-leucophoea* (1 of 1 trio). The only trios that were not significant were trios in which *azuay* was specified as P1 in the trio (i.e., more closely-related to one of the two taxa being measured for introgression).

We also used the program TreeMix to test for potential historical gene flow across subspecies. Using all five possible migration events specified in TreeMix, we found that the program indicated these top five patterns of migration (in order of strength from strongest to weakest): 1) *alticola* to *leucophoea* (migration edge weight 0.31), 2) *leucophoea* to *amazilia* (migration edge weight 0.27), 3) *azuay* to *amazilia* (migration edge weight 0.07), 4) *caeruleigularis* to *alticola* (migration edge weight 0.02), and 5) *caeruleigularis* to *dumerilii* (migration edge weight 0.01) (Fig 2.8). However, based on current geography and the strength of migration weight estimates, the only likely migration patterns specified by TreeMix are patterns 1 and 2 (see Fig. 2.7; migration pattern 1 matches up with scenario 4 in Fig. 2.7, and migration pattern 2 matches up with scenario 5 in Fig. 2.7), given *azuay* and *amazilia*, *caeruleigularis* and *alticola*, and *caeruleigularis* and *dumerilii*, respectively, are geographically distant from one another. Therefore, we only report the 3-population and 4-population statistical tests supporting the first two TreeMix migration patterns (i.e., 1) *alticola* to *leucophoea* and 2) *leucophoea* to *amazilia*).

Using 3-population tests implemented within the program TreeMix, we found support for migration pattern 2 from the subspecies *leucophoea* to *amazilia*, as one of

four 3-population tests with *amazilia* specified as the admixed population and with *leucophoea* included (i.e. [*amazilia*: *leucophoea*, x], where x = another subspecies) indicated admixture (f_3 = -0.0001, Z = -5.2, $p < 0.001$). We did not find support for migration pattern 1, as all four tests with *alticola* specified as the admixed population with *leucophoea* included (i.e. [*alticola*: *leucophoea*, x], where x = another subspecies) did not indicate admixture ($f_3 \geq 0.00007$; positive f_3 values indicate no admixture). With 4-population tests, the aim was to find 4-population trees that match the real-tree topology but that fail the 4-population test (i.e., have a |Z-score| $\geq$ 2), which would suggest the presence of gene flow across the true tree. We found support for migration pattern 1) *alticola* to *leucophoea* in three trees ($f_4 \geq 0.00007$, $Z \geq 5.1$, $p < 0.001$ for all three trees), and support for migration pattern 2) *leucophoea* to *amazilia* in one tree (f_4 = 0.0001, Z = 6.4, $p < 0.001$). Therefore, 3-population tests supported migration pattern 2 from *leucophoea* to *amazilia*, whereas 4-population tests supported migration from both migration pattern 1: *alticola* to *leucophoea* and migration pattern 2: *leucophoea* to *amazilia*.

Habitat Differentiation

To examine if environmental differences exist in the habitats of each subspecies, we assessed whether annual mean temperature and annual precipitation were significantly different across subspecies (given we only have one GPS point for *caeruleigularis*, all comparisons with *caeruleigualris* were non-significant). The average annual mean temperature ($\pm$ SE) in ºC for all subspecies were as follows: *alticola* (n = 7) 18.1 $\pm$ 1.3, *amazilia* (n = 12) 18.8 $\pm$ 0.2, *azuay* (n = 3) 15.9 $\pm$ 1.5, *caeruleigularis* (n = 1) 20.7 $\pm$ X,

dumerilii (n = 28) 24.3 ± 0.27, and *leucophoea* (n = 28) 19.2 ± 0.67. Using a Kruskal-Wallis test, temperature was significantly different across subspecies (Kruskal-Wallis, X^2 = 45.8, df = 5, p <0.001). Using post-hoc pairwise Dunn tests with a Bonferroni correction, *dumerilii* was significantly different from all other subspecies (except *caeruleigularis*) in annual mean temperature (Table 2.3).

The average annual precipitation (± SE) in mm for all subspecies was as follows: *alticola* (n = 7) 956.1 ± 109.9, *amazilia* (n = 12) 17.4 ± 5.9, *azuay* (n = 3) 616.7 ± 15.4, *caeruleigularis* (n = 1) 5 ± X, *dumerilii* (n = 28) 899.1 ± 97.8, and *leucophoea* (n = 28) 217.0 ± 50.2. Using a Kruskal-Wallis test, annual precipitation was significantly different across subspecies (Kruskal-Wallis, X^2 = 52.6, df = 5, p <0.001). Using post-hoc pairwise Dunn tests with a Bonferroni correction, there were significant differences between *alticola-amazilia*, *alticola-leucophoea, dumerilii-amazilia*, and *dumerilii-leucophoea* in annual precipitation (Table 2.3). To further visualize the differences in annual mean temperature and precipitation across subspecies, we plotted annual mean temperature vs. annual precipitation for each subspecies' GPS points (Fig. 2.9, 80% CI ellipses for each subspecies drawn for visualization).

For our second method of examining environmental differences across subspecies, we conducted a PCA on all 19 of the Worldclim2 variables from each GPS point and compared these values across subspecies. Overall, we found that the PCA was able to distinguish between subspecies using environmental variables (Fig. 2.10). We report only the values of the first five principal components (PCs), as they cumulatively explain 93.4% of the variance across subspecies. PC1 explained 40.3% of the variation, and was positively associated with mean diurnal range [0.12], temperature seasonality

[0.1], temperature annual range [0.19], and precipitation in the driest month [0.002], and negatively associated with all other Worldclim2 variables (Table 2.4). PC1 was significantly different across subspecies (Kruskal-Wallis test, $X^2 = 53.9$, df = 5, $p < 0.001$). Using post-hoc pairwise Dunn tests with a Bonferroni correction, PC1 was significantly different between *dumerilii-alticola* ($Z = 3.4$, $p = 0.005$), *dumerilii-amazilia* ($Z = 5.71$, $p < 0.001$), *dumerilii-azuay* ($Z = 3.44$, $p = 0.004$), and *dumerilii-leucophoea* ($Z = -5.88$, $p < 0.001$) (all other pairwise comparisons $|Z| < 2.06$, $p > 0.3$). Therefore, PC1 can be interpreted as distinguishing *dumerilii* from all other subspecies (Fig. 2.10).

PC2 explained 31.1% of the variation, and was positively associated with mean diurnal range [0.13], isothermality [0.34], annual precipitation [0.34], precipitation in the wettest month [0.28], precipitation in the driest month [0.29], precipitation in the wettest quarter [0.29], precipitation in the driest quarter [0.31], precipitation in the warmest quarter [0.20], and precipitation in the coldest quarter [0.23], and was negatively associated with all other Worldclim2 variables (Table 2.4). PC2 was significantly different across subspecies (Kruskal-Wallis test, $X^2 = 37.7$, df = 5, $p < 0.001$). Using post-hoc pairwise Dunn tests with a Bonferroni correction, PC2 was significantly different between *alticola-amazilia* ($Z = 4.91$, $p < 0.001$), *alticola-leucophoea* ($Z = 3.76$, $p = 0.001$), *amazilia-azuay* ($Z = -3.65$, $p = 0.002$), *amazilia-dumerilii* ($Z = -4.04$, $p = 0.004$), and very close to significant between *azuay-leucophoea* ($Z = 2.66$, $p = 0.06$) (all other pairwise comparisons $|Z| < 2.46$, $p > 0.10$). Therefore, PC2 can be seen as distinguishing *alticola*/*azuay* from *amazilia*/*leucophoea*, as well as additional differentiation between *amazilia* and *dumerilii* (Fig. 2.10).

PC3 explained 11.2% of the variation, and was positively associated with annual mean temperature [0.01], mean diurnal range [0.56], isothermality [0.27], max temperature warmest month [0.10], temperature annual range [0.40], mean temperature wettest quarter [0.12], mean temperature coldest quarter [0.2], annual precipitation [0.02], precipitation wettest month [0.15], precipitation seasonality [0.45], precipitation wettest quarter [0.01], and precipitation warmest quarter [0.06], and was negatively associated with all other Worldclim2 variables (Table 2.4). PC3 was significantly different across subspecies (Kruskal-Wallis test, $X^2 = 17.0$, df = 5, p = 0.004). Using post-hoc pairwise Dunn tests with a Bonferroni correction, PC3 was significantly different between *amazilia-leucophoea* (Z = -3.39, p = 0.005). Although PC4 explained 5.8% of the variation and was significantly different across subspecies (Kruskal-Wallis test, $X^2 = 13.1$, df = 5, p = 0.02), no subspecies comparisons were significant using the post-hoc pairwise Dunn tests with a Bonferroni correction (all comparisons p > 0.13). Similarly. PC5 explained 4.9% of the variation, but was not significantly different across subspecies (Kruskal-Wallis test, $X^2 = 6.4$, df = 5, p = 0.27).

Discussion

Phylogeny

We found that our phylogeny of 86 *A. amazilia* individuals based on our genomic dataset of 34,896 SNPs matches the current classification of the six subspecies. There were six clearly defined subspecies in the phylogeny with high bootstrap support, as all individuals within each subspecies clustered together, and no individuals were mismatched across subspecies. In addition, the subspecies fell into three distinct clades

associated with geography: a lowlands Ecuador clade (*dumerilii*), a highlands Ecuador clade (*alticola* and *azuay*), and a Peruvian clade (*leucophoea*, *amazilia*, and *caeruleigularis*). Although it has often been questioned whether the *A. amazilia* subspecies *alticola* (the Loja Hummingbird) should be considered a full species (Weller 2000, Krabbe and Ridgely 2010), our phylogeny suggests that *alticola* should remain a subspecies, due to its genetic similarity with *dumerilii* and *leucophoea*. In contrast, *azuay* is by far the most genetically distinct subspecies, and consideration could be taken into its species status, although it does it appear that *azuay* arose from within *alticola* lineage.

In opposition to the predictions made by Weller (2000) that *leucophoea* and *dumerilii* would form a northern clade due to phenotypic similarity, and that a south to north evolution of the subspecies was likely based on phenotypic characters, *dumerilii* is actually the most basal lineage of the phylogeny, with *leucophoea* more closely-related to *amazilia* and *caeruleigularis* than to *dumerilii*. This result is interesting because as Weller (2000) stated, both *leucophoea* and *dumerilii* have a prominent white throat patch and red underbelly, and are larger in size. However, phenotypic similarity between *leucophoea*, *dumerilii*, and *alticola* may stem from gene flow across these three subspecies (see below).

Gene Flow and Population Structuring Across Subspecies

In addition to the tree-building program RAxML, we used five different genomic analyses designed to assess gene flow and population structuring across the six subspecies of *A. amazilia*. Taking all methods into account, our results provide strong evidence of gene flow across the three subspecies of *dumerilii*, *leucophoea*, and *alticola*

(scenarios 1, 3, and 4 from Fig. 2.7), with support for introgression and/or strong genetic similarity from at least 5 of the 6 methods for each pair (see Table 2.5 for summary). These three subspecies have the potential for large areas of contact across the southern regions of Ecuador and the Ecuadorian-Peruvian border (Schulenberg et al. 2010, Weller et al. 2019), which could facilitate gene flow across areas of contact. In addition, both the tree and the PCA support introgression, as there are short branch lengths in the separation of *alticola*, *dumerilii*, and *leucophoea* in the tree compared to the other subspecies, and there is clear overlap in the genomic PCA of these three subspecies. fastStructure also supports genetic similarity, and in scenarios $K \geq 4$, two of the three subspecies were grouped together into one genetic population. Phenotypically, the three subspecies appear most similar to each other, with all three possessing a prominent white throat patch, a rufous underbelly, and a green or rufous backside and tail coverts. Phenotypic intergradation is visible between *alticola* and *dumerilii* in Ecuador, particularly in tail covert color and the amount of underside rufous belly coloration (Weller et al. 2019, S. Cowles personal observation).

In contrast, it appears that the subspecies *azuay* is the most genetically isolated subspecies with little to no gene flow from other subspecies, as it had the longest branch in the RAxML tree, the first principal component clearly separated *azuay* from the rest of the subspecies in the genomic PCA, and it was the most genetically differentiated from other subspecies in fastStructure for $K = 2$. The only method that supported any gene flow between *azuay* and another subspecies was Dsuite; however, Dsuite found gene flow in all subspecies pairwise scenarios (and may be untrustworthy at the subspecies level). Phenotypically, *azuay* is the only subspecies with little to no rufous coloration on

the chest and belly and an incomplete gorget (Krabbe and Ridgely 2010, S. Cowles personal observation). *Caeruleigularis* and *amazilia* were the next most genetically isolated subspecies, based on the RAxML tree branch lengths, the second principal component in the genomic PCA, and the fastStructure results for K=3. *Amazilia* and *caeruleigularis* are the only subspecies to lack both a white throat patch and white belly coloration (as the belly is entirely rufous), and have distinctly colored gorgets (dark green and bluish-purple, respectively) compared to the medium green of the other four subspecies (Weller 2000). In addition, our results suggest that there was some introgression between *leucophoea* and *amazilia* in Peru (Scenario 5 from Fig. 2.7), particularly supported by our PCA, TreeMix, and fastStructure results.

Although our different genomic methods gave slightly contrasting results, all methods together gave an overall conclusion of gene flow between *alticola*, *dumerilii*, and *leucophoea*, some gene flow between *leucophoea* and *amazilia*, and strong genetic isolation of *azuay* and *caeruleigularis*. We suggest that it is imperative to use multiple genomic methods together to obtain a clear overall picture of gene flow between highly closely-related lineages, such as taxa at the subspecies level. Some methods are not designed to or able to detect gene flow in extremely closely-related taxa like subspecies and may have better potential for assessing species-level relationships. For example, the fineRADstructure clustering algorithm may work better for more highly-differentiated populations at the species level, as it was unable to detect any population structuring beyond the two most genetically differentiated subspecies *azuay* and *caeruleigularis*. Similarly, the Dsuite algorithm may work better for testing introgression at the species level, as it detected introgression in all five introgression scenarios tested (*alticola*-

dumerilii, *azuay-dumerilii*, *dumerilii-leucophoea*, *alticola-leucophoea*, and *amazilia-leucophoea*). The only trios that were not significant were trios in which *azuay* was specified as P1 in the trio. This may be due to the fact that Dsuite and the calculation of D-statistics only take into account tree topology and not branch length. We know that *azuay* is the most genetically distinct subspecies by far from both the branch lengths of the phylogeny and the PCA, and not taking branch lengths into account likely affected the calculation of the D-statistics.

Even though there is support for introgression between *azuay* and *dumerilii* using D-statistics (scenario 2), the phylogenetic tree and PCA suggest otherwise. In both of these *azuay-dumerilii* trios, P1 was either *amazilia* or *caeruleigularis*, which means the three subspecies used in the trio had extremely long branch lengths and none were particularly closely-related to one another. Given that there is a signature of latitudinal differentiation in the genomic structure (PC2 in the PCA separates subspecies north to south across their ranges), it would make sense that D-statistics would predict that a mid-latitude subspecies (i.e. *dumerilii*) shows introgression with the most northern subspecies (*azuay*) compared to the most southern subspecies (*amazilia* or *caeruleigularis*) when the most southern subspecies (*amazilia* or *caeruleigularis*) is supposed to be more closely related to the northern subspecies (*azuay*) based on the trio topology. Therefore, we suggest that *alticola*, *dumerilii*, and *leucophoea* have a history of introgression, as well as *leucophoea-amazilia*, but that *azuay* and *dumerilii* do not.

We also found it interesting that we did not find more well-supported within-subspecies clustering, especially in our *dumerilii* field-collected samples from Ecuador. We collected *dumerilii* samples from field sites over 200 kilometers apart, and even

separated by a water barrier (Ayampe, Guayaquil, and Isla Puná), yet did not recover any well-supported phylogenetic structuring (bootstrap values ≥70) for these groups. This result suggests potentially high levels of gene flow and movement within each subspecies range.

mtDNA patterns and timing of divergence

We built a haplotype network based on 1002bp of the mitochondrial gene *ND2*. From the network, we can conclude that five of the six subspecies share a common haplotype for this gene (all but *caeruleigularis*), which is in contrast to the phylogeny based on genomic data that shows clear subspecies groupings. Most surprisingly, the majority of *azuay* individuals share this common species haplotype, even though the genomic data from nuclear SNPs placed *azuay* as the most genetically isolated subspecies. *Caeruleigularis* is the only subspecies with a differentiated haplotype for *ND2*, and is separated by approximately 1% sequence divergence, which equates to about 500,000 years using the standard mtDNA divergence rate for birds. For most of the subspecies, our 1002bp of the mtDNA *ND2* gene may not provide enough of a phylogenetic signal (i.e. enough variation) to recover subspecies relationships and timing of divergence. To obtain a better tree of the mitochondrial history and timing of divergence across subspecies, sequencing of additional mtDNA genes would be warranted.

However, in birds, it is common to find mismatches between histories based on mtDNA genes and genomic-wide data within phylogeographic studes (so called mito-nuclear discordance, see Funk and Omland (2003), McKay and Zink (2010), Toews and

Brelsford (2012), and Hill (2017) for comprehensive overviews), and this scenario can be indicative of isolation coupled with periods of gene flow (e.g. Webb et al. 2011, Hogner et al. 2012, Block et al. 2015, Zarza et al. 2016), selection on specific mtDNA genes (e.g. Morales et al. 2015), differential introgression of mtDNA versus nuclear DNA (e.g. Carling & Brumfield 2008, Sardell and Uy 2016), or incomplete lineage sorting of mtDNA haplotyptes (e.g. Harvey and Brumfield 2015). In the congeneric species *A. tzacatl* that ranges from Southeastern Mexico down to the central coast of Ecuador, Miller et al. (2011) found five distinct mtDNA clades with divergence up to 2.4%; however, these clades did not fully correspond with the five currently recognized subspecies of *A. tzacatl*. In our case with *A. amazilia*, we found that although genomic data show that the six subspecies are genomically distinct from one another, the mtDNA haplotypes for each subspecies are not distinct from one another except for the case of *caeruleigularis*. This pattern suggests that assessing taxonomic histories based on small amounts (i.e., one or two mtDNA genes) of mtDNA alone, particularly within the case of closely-related taxa like subspecies, should be done with caution, and that genomic-wide datasets in concert with mtDNA are more likely to provide the necessary level of resolution.

Given that *caeruleigularis* evolved within the last 500,000 years and is one of the later nodes in the phylogenetic tree, we could expect that the timing of divergence for other subspecies groupings would fall within the range of 3 million to 500,000 years ago (during the Pleistocene), given *A. amazilia* and one of its closet congeners *A. franciae* (see Ornelas et al. 2014) are estimated to have been separated for approximately 3.25 million years based on our *ND2* data. Our estimates of divergence times between *A.*

amazilia and its two congeners *A. franciae* and *A. tzacatl* differ according to whether we use the standard 2% divergence per million years in mtDNA genes as is commonly done for birds (Price 2008, Cowles and Uy 2019) or use the unconventional clock rate of 0.0068 base substitutions per million years as in Ornelas et al. (2014). McGuire et al. (2015) also found discrepancies with the timing predictions made by Ornelas et al. (2013) in species accumulation based on the Panamanian uplift, and our estimates of divergence times between *A. amazlia* and its congeners *A. franciae* and *A. tzacatl* are better in line with McGuire et al. (2015)'s later estimates that most of the Amazilia genus evolved in the late Miocene and Pliocene (11.63-2.58 million years ago).

Expansion followed by geographic isolation

According to our two analyses assessing temperature and precipitation variables, the six subspecies of *A. amazilia* inhabit three fairly distinct habitat types: hotter and wetter (subspecies *dumerilii*, tropical dry forest in the Ecuador lowlands), cooler and wetter (subspecies *azuay* and *alticola*, subtropical forest in the Ecuador Highlands), and drier and intermediate in temperature (subspecies *leucophoea*, *amazilia*, and *caeruleigularis*, desert scrub and dry coastal environments on the Peruvian coast). These distinct habitats suggest that ecological expansion into new habitats from the Ecuador lowlands into the highlands and into the desert environments of Peru was an important driver of subspecies diversification in *A. amazilia*. Subsequently, the isolation of *azuay* from *alticola* in the Yunguilla Valley of Ecuador, and *amazilia* and *caeruleigularis* from *leucophoea* in the Rio Pativilca and Nazca/Ica regions of Peru, respectively, likely drove additional subspecies differentiation. We did not find any significant differences in

environmental characteristics between *alticola* and *azuay*, and very few differences between *leucophoea*, *amazilia*, and *caeruleigularis* (only PC3 was different between *leucophoea* and *amazilia* in the PCA of the 19 Worldclim variables; temperature and precipitation were not significantly different between the two), supporting at least a partial role for geographical isolation in driving the divergence of *azuay* from *alticola*, and *amazilia* and *caeruleigularis* from *leucophoea*. Our results suggest that a combination of both ecological expansion into new habitats and geographical isolation were critical drivers of subspecies diversification in *A. amazilia*. These results are consistent with Graham et al. (2009)'s broader study of hummingbird communities, which found that temperature, precipitation, and vegetation structure were important factors in structuring hummingbird community composition.

The subspecies *azuay* is found only within the Yunguilla Valley (Río Jubones drainage system) of the Azuay province of Ecuador, which is a 104-mile dry valley that extends from south of Cuenca to Machala along the coast. This valley is completely isolated on either side, given its border with the humid tropical forest of the western Andes on the western side, and the humid tropical forest of Cordillera Chilla-Tioloma-Fierrourcu and the Azuay-Morona-Zamora ranges on the eastern side (Krabbe and Ridgely 2010). Another bird species, the Pale-headed Brush Finch (*Atlapetes pallidiceps*) is endemic to the Yunguilla Valley (Krabbe and Ridgely 2010), which shows that this valley has been an isolated refuge for some time and allowed for the divergence of the *azuay* subspecies.

Similarly, the most southern subspecies *caeruleigularis* is also found within a well-isolated valley: the small Ica-Nazca region of Peru. This region is surrounded by dry

desert plains; however, the large Pasco and Ica Rivers in the Rio Grande de Nazca drainage have provided a water source in the dry desert conditions since ancient times. Within the Ica and Nazca River Valleys, other rare Peruvian endemics, such as the plant *Prosopis limensis* and *Onoseris humboldtiana*, are found (Whaley et al. 2019). Interestingly, the Nazca Valley has been home to one of the most ancient human civilizations due to its water source. The Nazca people developed subterranean irrigation systems (Schreiber and Lancho Rojas 1995), and today, human-induced agriculture in the Nazca Valley might have had positive impacts on *caeruleigularis* through the increase of food availability.

Finally, we note that the most geographically isolated populations of *azuay*, *amazilia*, and *caeruleigularis* are the most genetically distinct subspecies. This could be due to genetic bottlenecks that occurred during the initial isolation of a few individuals in the formation of these subspecies, which allowed for genetic drift and small population size to play an important role in genetic divergence. Woolfit and Bromham (2005) demonstrated a trend of higher nonsynonymous substitution rates in a variety of taxa comparing substitution rates in small populations isolated on islands with their closely-related mainland taxa. Particularly in birds, Johnson and Seger (2001) found that ducks and doves on small islands had higher nonsynonymous substitution rates (but no difference in synonymous substitution rates) in mtDNA genes compared to their closely-related mainland relatives, and Smith and Klicka (2013) found that mtDNA substitution rates were increased on island populations of cardinals. However, it is unknown whether this trend is well-supported in the nuclear DNA of birds, or in small isolated mainland populations. In contrast, increased genetic differentiation could be due to the rapid

adaptation of an isolated population to novel habitat conditions. However, given that the environmental conditions seem to be similar between *alticola* and *azuay*, and *amazilia*, *caeruleigularis*, and *leucophoea*, this hypothesis seems less plausible.

Final Summary

The divergence of the six subspecies of *A. amazilia* has likely been shaped by both range expansion and geographical isolation. Both genetic and geographic patterns are consistent with an expansion of the prototype *A. amazilia* from the lowlands of Ecuador into the highlands of Ecuador and the dry environments of Peruvian coast, followed by secondary divergence of *azuay*, *amazilia*, and *caeruleigularis* due to geographic isolation. Of the six subspecies, *azuay* is the genetically the most distinct, and seems to have arisen from within the *alticola* subspecies in the Yunguilla Valley of Ecuador. Genetic evidence points to past introgression and gene flow across the three subspecies of *alticola*, *dumerilii*, and *leucophoea*. Further research should explore the evolutionary history of multiple species with wide geographical ranges, as studying the early stages of speciation provides insight into the factors underlying the initial stages of diversification.

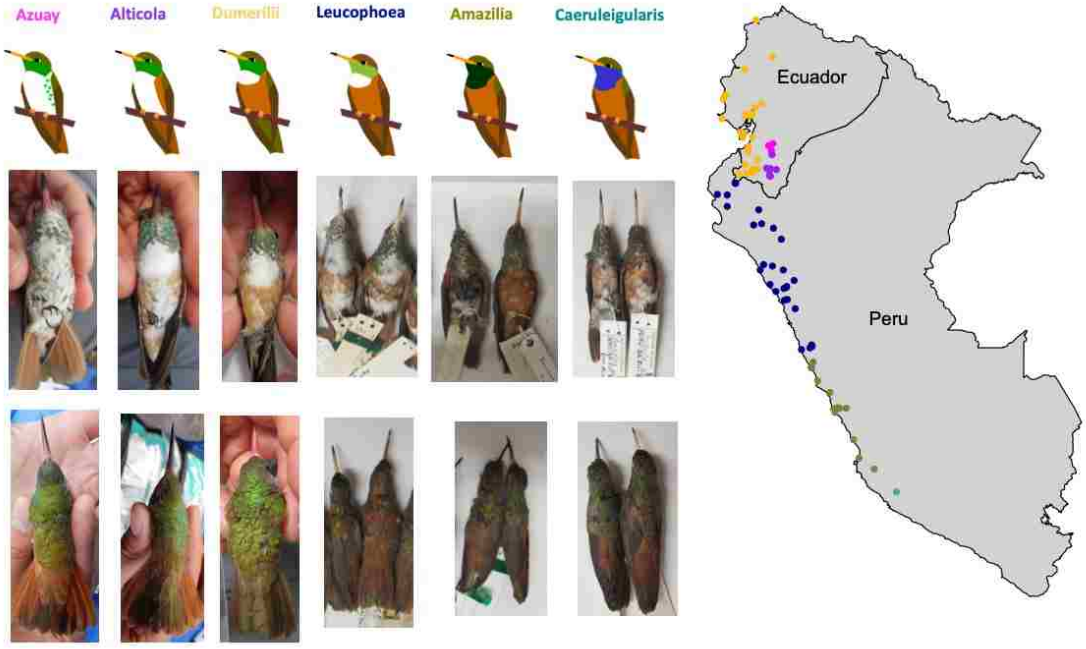

Figure 2.1: Range of the six subspecies of *A. amazilia* across Ecuador and Peru. Points on map indicate historical collection records (from Weller 2000) as well as our field collection sites (see Table 2.1).

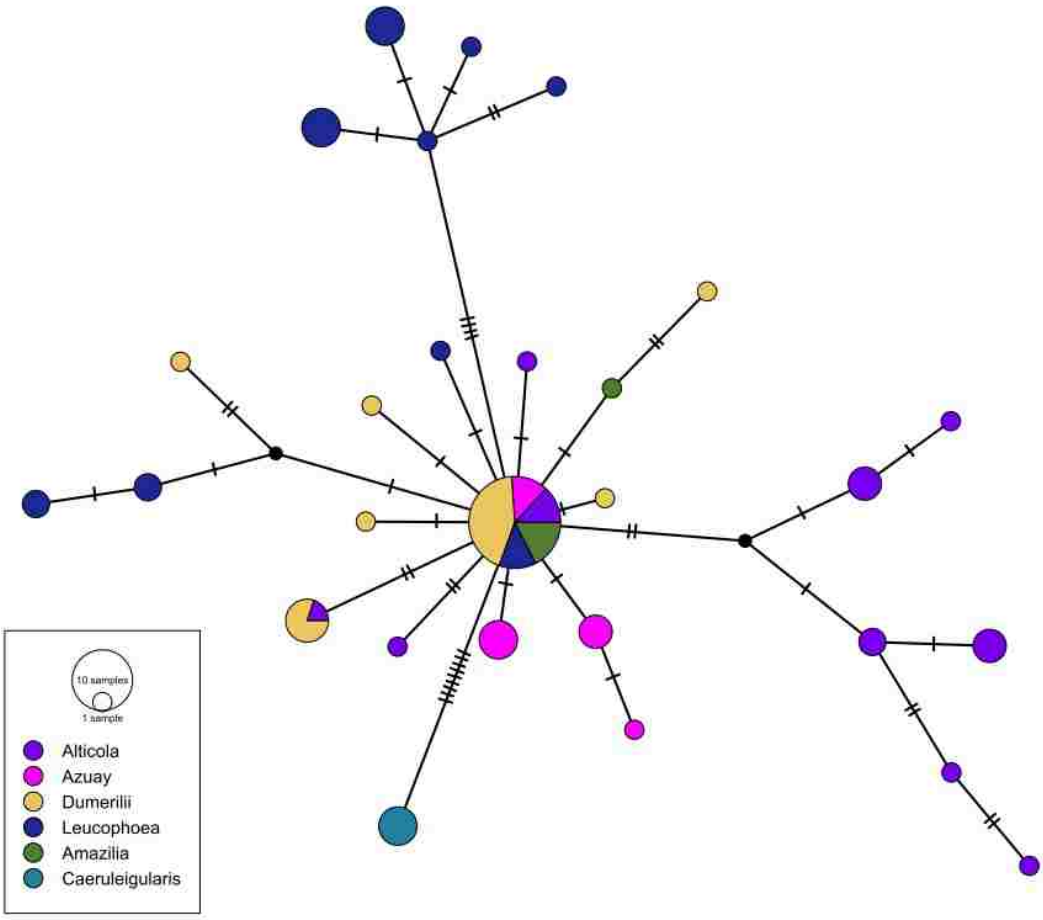

Figure 2.2: *ND2* haplotype network for *A. amazilia*. The median-spanning haplotype network for 75 individuals across subspecies is based on 1002 bp of the mtDNA gene *ND2*. *Caeruleigularis* is the only subspecies that has a distinct haplotype network. All other subspecies share a common species haplotype (23 individuals share the common haplotype across subspecies).

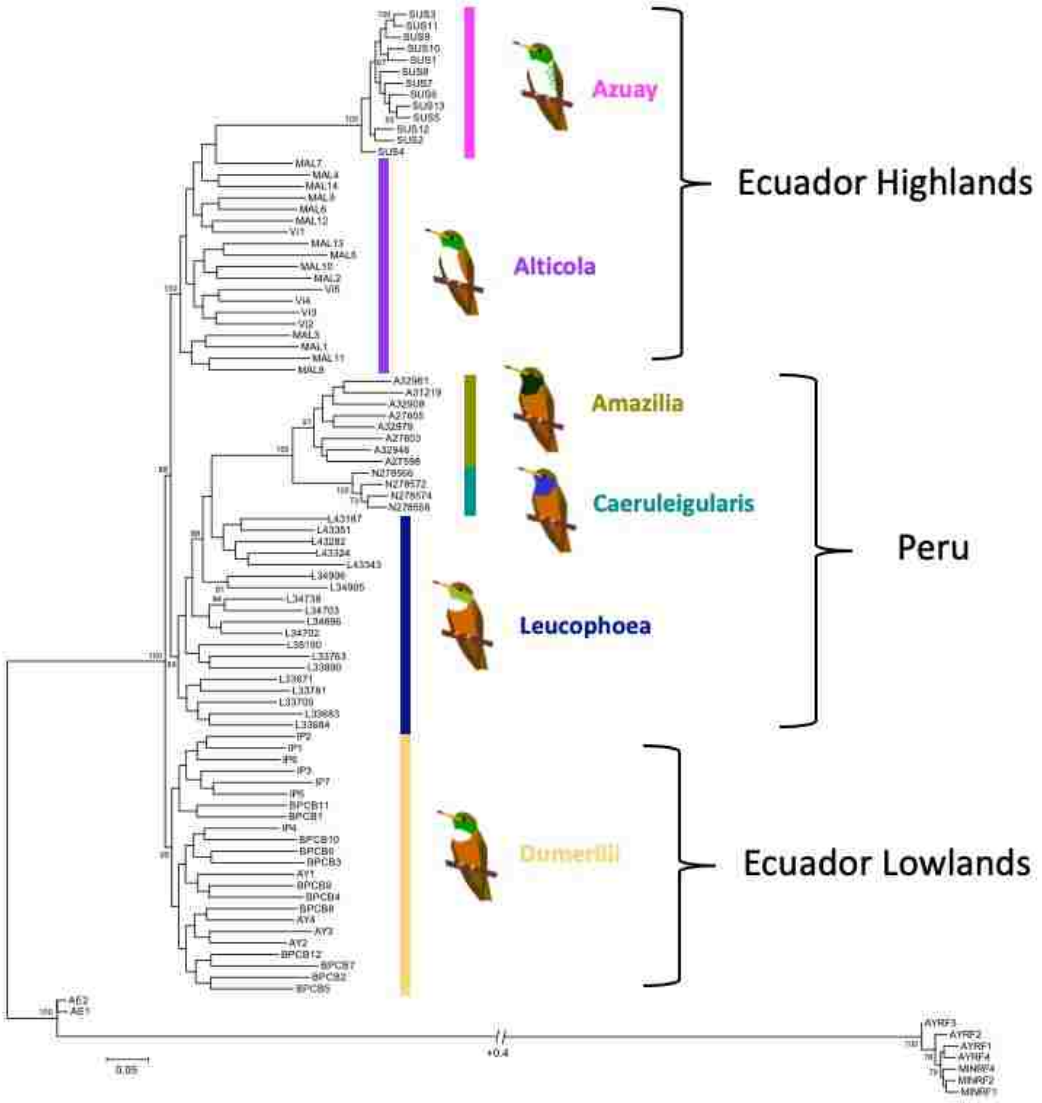

Figure 2.3: RAxML phylogenetic tree of the *A. amazilia* subspecies complex. Our maximum likelihood-based tree from the program RAxML is based on our dataset of 34,896 SNPs for all six subspecies of *A. amazilia* (86 individuals total) and 9 congener samples (2 *A. franciae* and 7 *A. tzacatl*). Only bootstrap percentages $\geq$ 70 are shown on the tree.

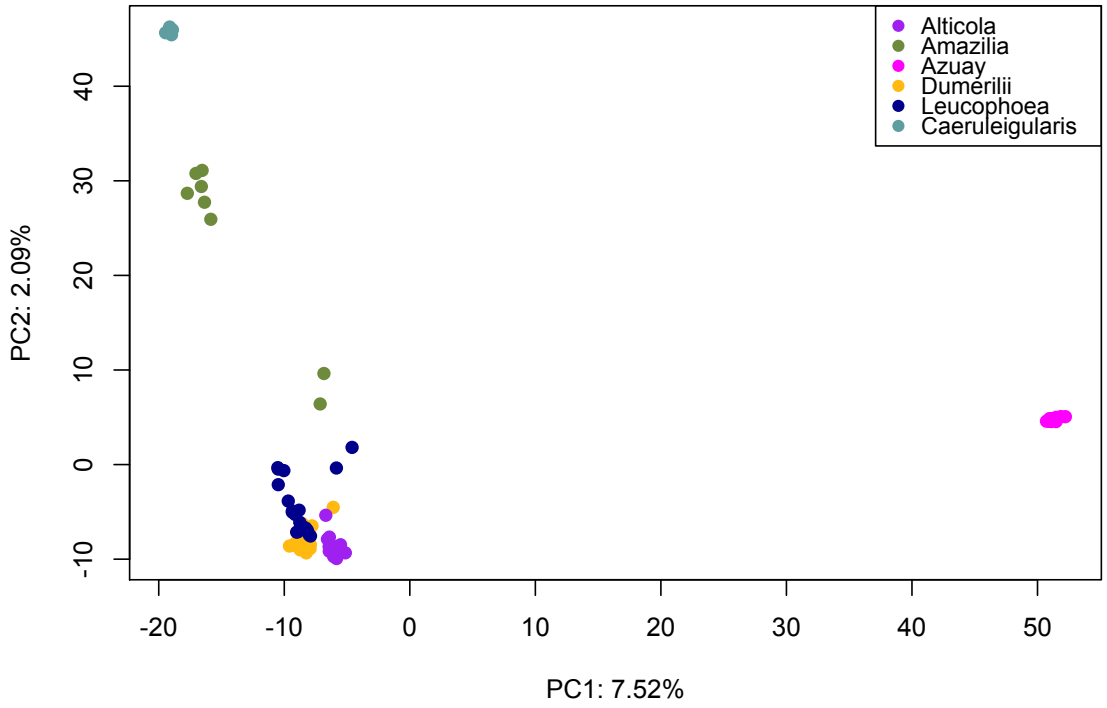

Figure 2.4: PCA of genomic data. We conducted a PCA on the genomic dataset of 34,896 SNPs in TASSEL using the 'LD-kNNi' impute method for missing data (see text for details). PC1 separates *azuay* from the other five subspecies, whereas PC2 separates *caeruleigularis* and most *amazilia* individuals from a cluster of *alticola*, *dumerilii*, and *leucophoea*.

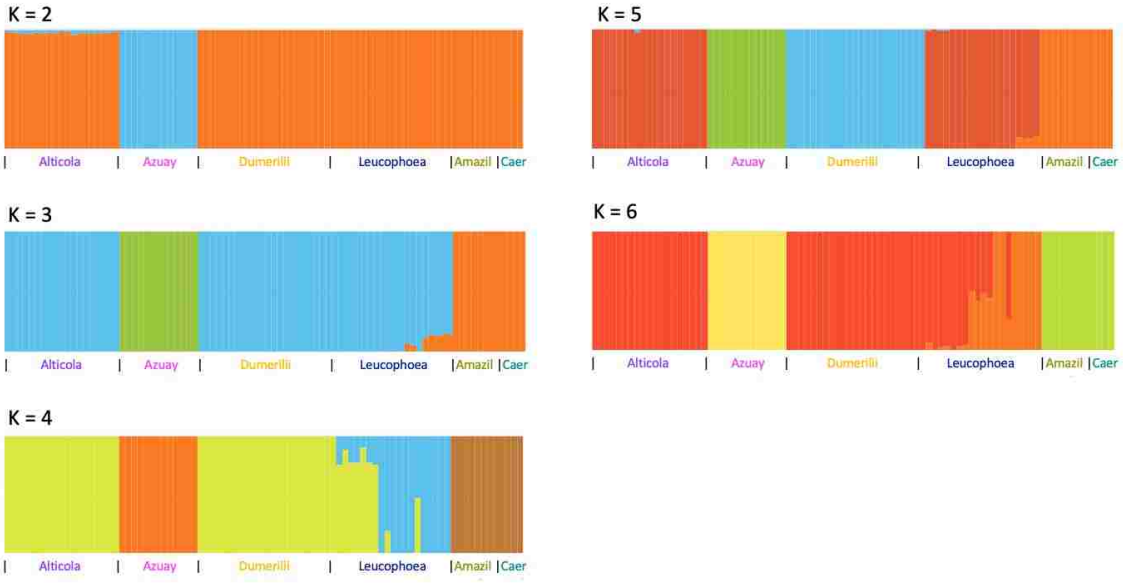

Figure 2.5: fastStructure plots for K=2-6. K = 4 was the best-fit model using the choose.py script for K $\geq$ 4. Different colors on plots represent different assigned genetic clusters.

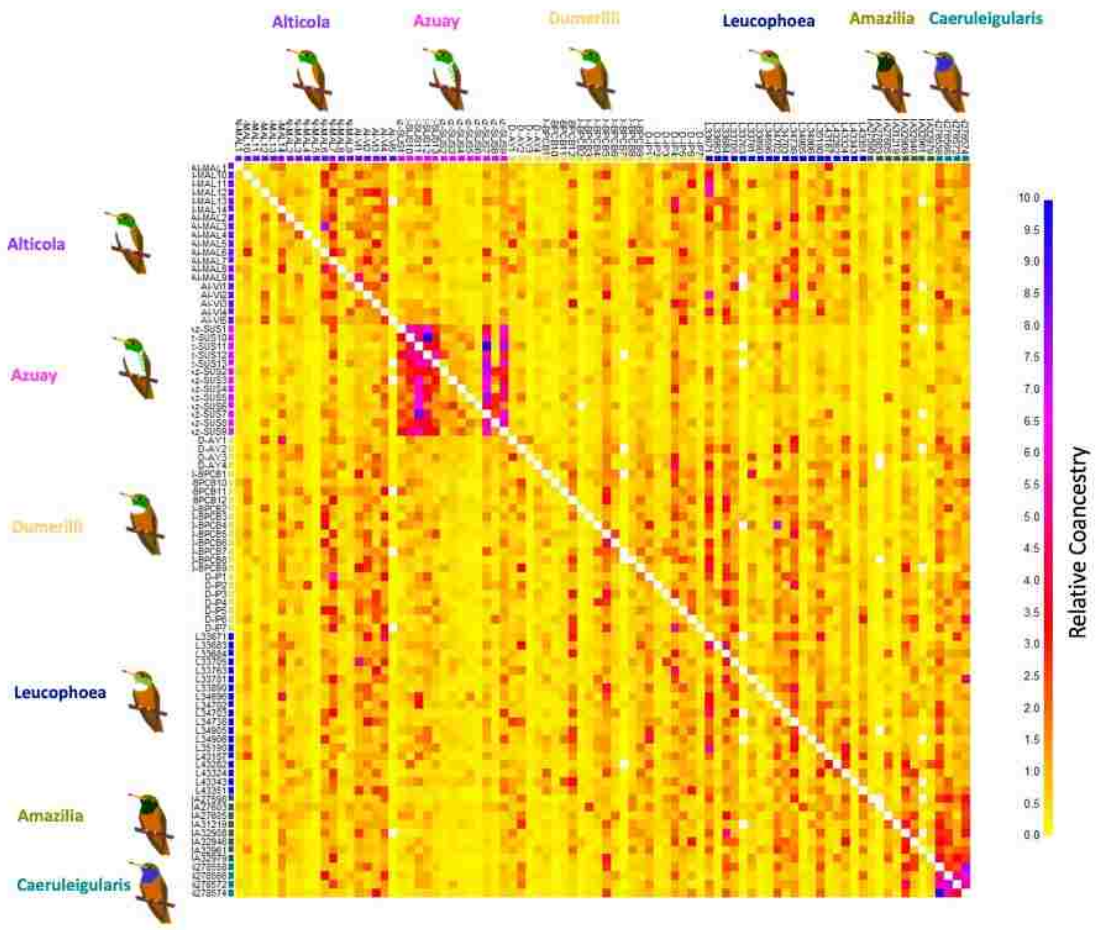

Figure 2.6: fineRADstructure coancestry pairwise plot for individuals organized by subspecies. Colors on heatmap represent strength of relative coancestry (blue = high, yellow = low) between pairs of individuals. Colors next to sample names identify subspecies. Only individuals from *azuay* and *caeruleigularis* showed elevated coancestry levels and subspecies clustering.

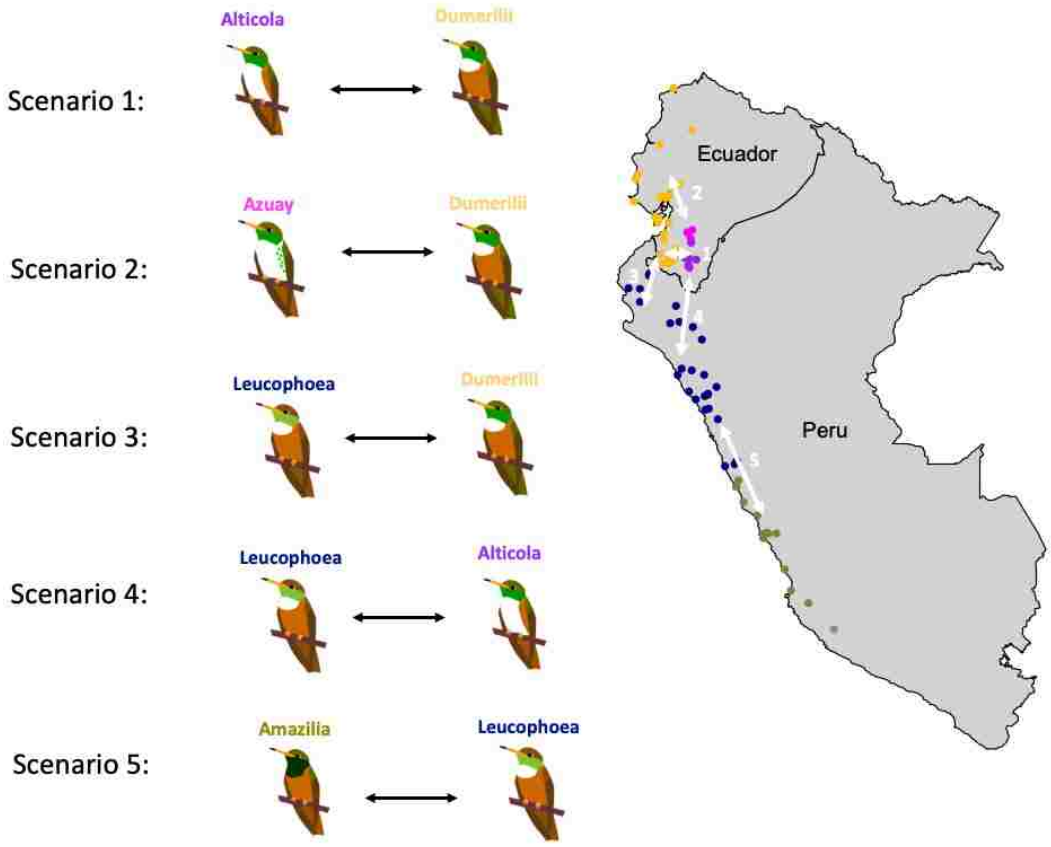

Figure 2.7: Five possible gene flow/introgression scenarios based on geography and the phylogenetic tree (pairs could not be sister taxa or derived from one another).

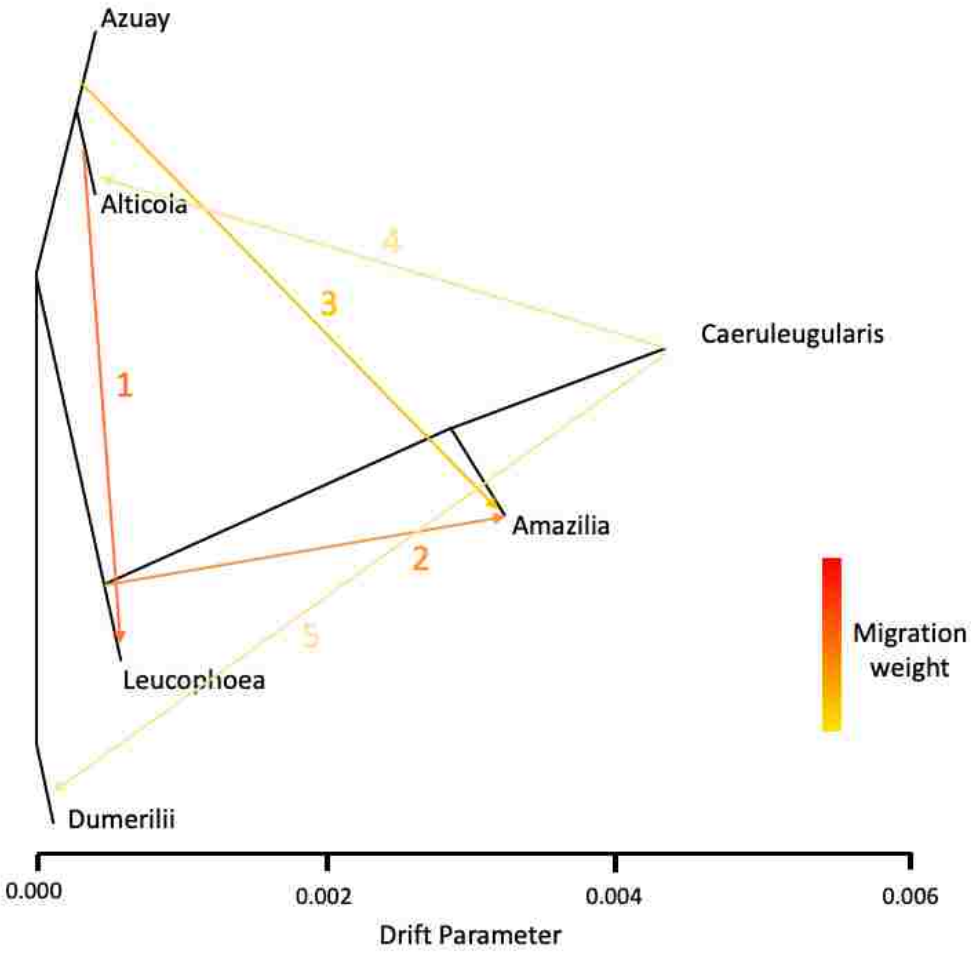

Figure 2.8: TreeMix results for five possible migration events. We used the RAxML tree as an input file to specify tree topology (tree is outlined in black). Colored arrows represent migration weight (red = strongest, yellow = weakest), and numbers represent ordinal number of the most likely migration events (i.e. 1-5) specified by the TreeMix program. Only scenarios 1) *alticola* to *leucophoea* and 2) *leucophoea* to *amazilia* have high migration edge weights (0.37 and 0.29, respectively; all other scenarios $\leq$ 0.07) and are in concordance with geography (i.e. subspecies that could contact one another). 3- and 4-population statistical tests also lend support to both scenarios 1 and 2 (see text for details).

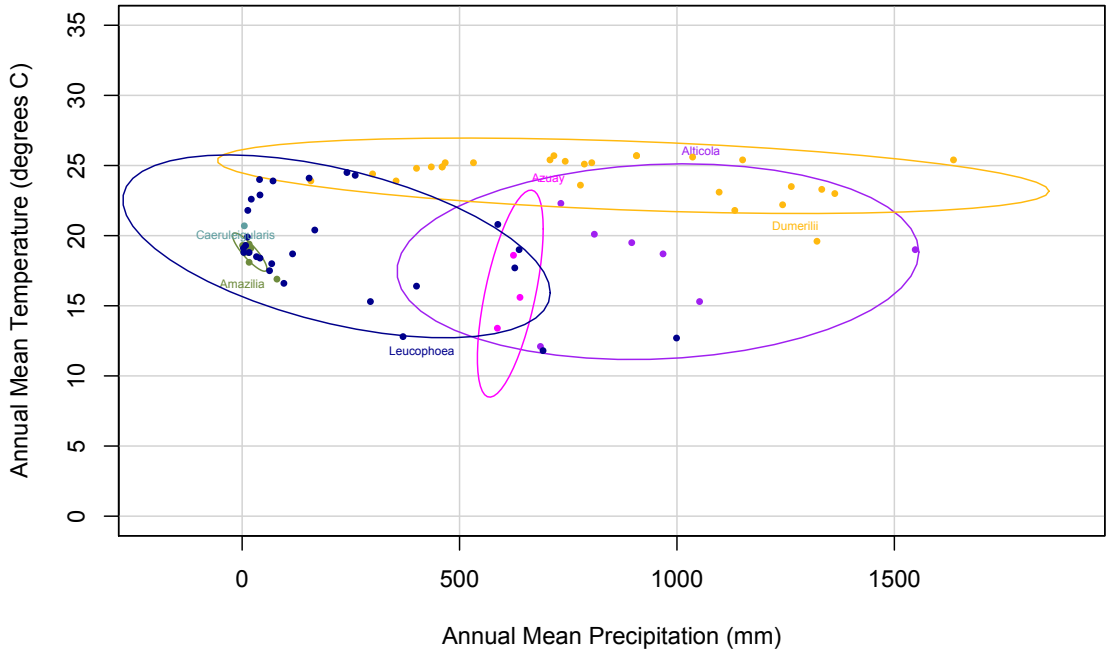

Figure 2.9: Annual mean temperature in ºC vs. annual precipitation in mm for each of the subspecies of *A. amazilia*. Ellipses are drawn at 80% confidence intervals for visualization.

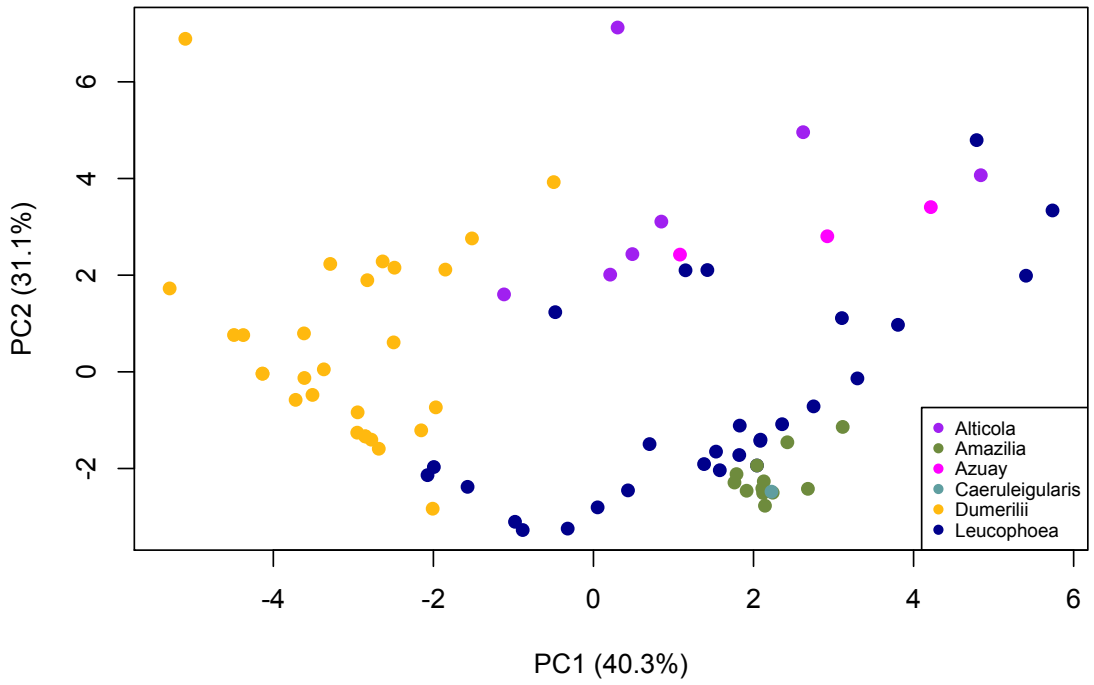

Figure 2.10: PC1 vs PC2 for all 19 Worldclim2 variables. Both PC1 and PC2 have significantly different subspecies groupings from one another (see text).

Table 2.1: Details of Ecuador and Peru Samples

ECUADOR	Species	Subspecies	Location	GPS (approx)	Elevation	# of Samples	MSB IDs
	Amazilia amazilia	Azuay	Susudel, Azuay	S3° 24.289' W79° 11.095'	2385	13	NA
		Alticola	Vilcabamba, Loja	S4° 15.653' W79° 13.204'	1482	5	NA
			Malacatos, Loja	S4° 11.691' W79° 14.717'	1616	14	NA
		Dumerilii	Bosque Protector Cerro Blanco, Guayas	S2° 10.836' W80° 01.216'	50	12	NA
			Cauchiche, Isla Puná, Guayas	S2° 47.595' W80° 14.300'	2	7	NA
			Ayampe, Manabí	S1° 40.695' W80° 48.741'	2	4	NA
					Ecuador *A. amazilia*	**55**	
	Amazilia tzacatl	jucunda	Ayampe	S1° 40.695' W80° 48.741'	2	4	NA
			Mindo	S0° 03.287' W78° 46.656'	1214	5	NA
	Amazilia franciae		Mindo	S0° 04.734' W78° 46.130'	1345	2	NA
					TOTAL CONGENERS:	11	

PERU	Amazilia amazilia	Subspecies	Location	GPS (approx)	Elevation	# of Samples	MSB IDs						
		Leucophoea	Olmos, Lambayeque	S5° 53.497' W79° 47.215'	126	7	33671	33683	33684	33705	33763	33781	33890
			Huacapongo, Viru, La Libertad	S8° 23.211' W78° 38.711'	309	5	34693	34696	34702	34703	34738		
			Rio Santa, Viru, La Libertad	S8° 41.691' W78° 22.717'	357	2	34905	34906					
			Bosque Cachil, Contumazá, Cajamarca	S7° 23.882' W78° 46.969'	2500	1	35190						
			Coris, Aija, Ancash	S9° 54.484' W77° 49.282'	1029	2	43187	43282					
			Yanaparin, Huarmey, Ancash	S9° 57.518' W77° 48.442'	723	2	43305	43324					
			Mal Paso, Huarmey, Ancash	S10° 0.451' W77° 52.896'	373	2	43343	43351					
		Amazilia	Huachipa, Lima, Lima	S12° 0.504' W76° 55.261'	366	7	27598	27603	27605	31219	32946	32961	32979
			Sisicaya, Huarochiri, Lima	S12° 1.77' W76° 39.047'	935	2	32901	32908					
		Caerugularis	Nazca, Ica	NA	~500	4	27598	27603	27605	31219			
					Peru *A. amazilia*	**34**							
					TOTAL *A. amazilia*	**89**							

Table 2.2: D-statistics and Bonferroni-corrected P-values for geographically possible ABBA-BABA subspecies trios (see Fig 2.7 for scenarios).

Scenario	P1	P2	P3	D-statistic	p-val (bonferroni corrected)
1	*amazilia*	*alticola*	*dumerilii*	0.077	<0.0001
1	*azuay*	*alticola*	*dumerilii*	0.019	0.514
1	*caeruleigularis*	*alticola*	*dumerilii*	0.089	<0.0001
1	*leucophoea*	*alticola*	*dumerilii*	0.015	0.023
2	*amazilia*	*azuay*	*dumerilii*	0.06	0.00025
2	*caeruleigularis*	*azuay*	*dumerilii*	0.072	<0.0001
3	*amazilia*	*leucophoea*	*dumerilii*	0.067	<0.0001
3	*azuay*	*leucophoea*	*dumerilii*	0.003	1
3	*caeruleigularis*	*leucophoea*	*dumerilii*	0.079	<0.0001
4	*amazilia*	*leucophoea*	*alticola*	0.087	<0.0001
4	*azuay*	*alticola*	*leucophoea*	0.015	1
4	*caeruleigularis*	*leucophoea*	*alticola*	0.101	<0.0001
5	*caeruleigularis*	*amazilia*	*leucophoea*	0.058	<0.0001

Table 2.3: Dunn test statistics for pairwise comparisons between subspecies using annual mean temperature and annual precipitation. All p-values have been Bonferroni-corrected for multiple comparisons.

TEMPERATURE	***alticola***	***amazilia***	***azuay***	***caeruleigularis***	***dumerilii***
amazilia	Z = -0.08 P = 1				
azuay	Z = 0.97 P = 1	Z = 1.1 P = 1			
caeruleigularis	Z = -0.67 P = 1	Z = -0.65 P = 1	Z = -1.2 P = 1		
dumerilii	Z = -3.48 **P < 0.001**	Z = -4.6 **P < 0.001**	Z = -3.78 **P = 0.0012**	Z = -0.9 P = 1	
leucophoea	Z = -0.4 P = 1	Z = -0.38 P = 1	Z = -1.38 P = 1	Z = 0.53 P = 1	Z = 5.45 **P < 0.001**
PRECIPITATION	***alticola***	***amazilia***	***azuay***	***caeruleigularis***	***dumerilii***
amazilia	Z = 4.67 **P < 0.001**				
azuay	Z = 0.93 P = 1	Z = -2.45 P = 0.11			
caeruleigularis	Z = 2.36 P = 0.14	Z = 0.30 P = 1	Z = 1.64 P = 0.76		
dumerilii	Z = 0.48 P = 1	Z = -5.84 **P < 0.001**	Z = -0.72 P = 1	Z = -2.28 P = 0.17	
leucophoea	Z = 3.58 **P = 0.003**	Z = -2.04 P = 0.31	Z = 1.44 P = 1	Z = -1 P = 1	Z = 4.9 **P < 0.001**

Table 2.4: PCA loadings for all 19 Worldclim variables for PC1-PC5.

Worldclim2 Variable	PC1	PC2	PC3	PC4	PC5
Annual Mean Temp	-0.34	-0.12	0.01	0.09	0.04
Mean Diurnal Range	0.12	0.13	0.56	0.27	0.002
Isothermality	-0.06	0.34	0.27	-0.08	0.18
Temp Seasonality	0.10	-0.35	-0.09	0.20	-0.23
Max Temp Warmest Month	-0.29	-0.21	0.10	0.27	0.01
Min Temp Coolest Month	-0.34	-0.08	-0.16	-0.03	0.10
Temperature Annual Range	0.19	-0.15	0.40	0.40	-0.15
Mean Temp Wettest Quarter	-0.31	-0.14	0.12	0.13	-0.04
Mean Temp Driest Quarter	-0.31	-0.13	-0.12	0.09	0.11
Mean Temp Warmest Quarter	-0.3	-0.22	-0.02	0.15	-0.03
Mean Temp Coldest Quarter	-0.35	-0.04	0.02	0.04	0.09
Annual Precipitation	-0.18	0.34	0.02	0.02	-0.07
Precipitation Wettest Month	-0.23	0.28	0.15	-0.12	-0.07
Precipitation Driest Month	0.002	0.29	-0.27	0.47	-0.26
Precipitation Seasonality	-0.16	-0.12	0.45	-0.04	-0.11
Precipitation Wettest Quarter	-0.23	0.29	0.10	-0.12	-0.06
Precipitation Driest Quarter	-0.005	0.31	-0.26	0.44	-0.25
Precipitation Warmest Quarter	-0.22	0.20	0.06	-0.20	-0.51
Precipitation Coldest Quarter	-0.3	0.23	-0.03	0.33	0.66

Table 2.5: Summary of evidence of introgression across five possible geographic scenarios (Fig. 2.7) using six different genomic analysis methods designed to work with RADseq data

Scenario	Subspecies 1	Subspecies 2	RAxML tree	PCA	faststructure	fineRAD structure	Dsuite	Tree Mix
1	*alticola*	*dumerilii*	Yes	Yes	Yes	Yes	Yes	
2	*azuay*	*dumerilii*					Yes	
3	*dumerilii*	*leucophoea*	Yes	Yes	Yes	Yes	Yes	
4	*alticola*	*leucophoea*	Yes	Yes	Yes	Yes	Yes	Yes
5	*leucophoea*	*amazilia*		Yes	Yes	Yes	Yes	Yes

Chapter 3

Rapid, complete reproductive isolation in two closely-related *Zosterops* White-eye bird species despite broadly overlapping ranges[1]

Summary

Examining what happens when two closely-related species come into secondary contact provides insight into the later stages of the speciation process. The Zosteropidae family of birds is one of the most rapidly speciating vertebrate lineages. Members of this family are highly vagile and geographically widespread, raising the question of how divergence can occur if populations can easily come into secondary contact. On the small island of Kolombangara, two closely-related non-sister species of White-eyes, *Zosterops kulambangrae* and *Z. murphyi*, are distributed along an elevational gradient and come into secondary contact at mid-elevations. We captured 134 individuals of both species along two elevational transects. Using genotyping-by-sequencing data and a mitochondrial marker, we found no evidence of past hybridization events and strong persistence of species boundaries, even though the species have only been diverging for approximately two million years. We explore potential reproductive barriers that allow the two species to coexist in sympatry, including premating isolation based on divergence in plumage and song. We also conducted a literature review to determine the time it takes to evolve complete reproductive isolation in congeneric avian species/subspecies in secondary contact (restricted to cases where congeneric taxa are parapatric or have a hybrid zone), finding that our study is one of the youngest examples of complete reproductive isolation studied in a genomic context reported in birds.

[1]This chapter has already been published as Cowles, S.A. and Uy, J.A.C. 2019. Rapid, complete reproductive isolation in two *Zosterops* White-eye bird species despite broadly overlapping ranges. Evolution 73:1647-1662. https://doi.org/10.1111/evo.13797

Background

Speciation can be a long and complex process, sometimes involving long periods of isolation followed by secondary contact (Dobzhansky 1937, Mayr 1942). If secondary contact occurs after a period of divergence in allopatry, several scenarios can occur (Mayr 1942, Liou and Price 1994, Coyne and Orr 2004). First, if the two taxa have diverged sufficiently in traits used for mate recognition or if hybrids are at a strong ecological or sexual disadvantage, the two taxa can remain reproductively isolated with little to no gene flow and exist in sympatry (depending on ecological differentiation and competition for resources, see Freeman 2015, Cooney et al. 2017). Second, if the taxa have not diverged in mate recognition systems or if hybrids do not have strong disadvantages and little to no reproductive isolation has evolved, the taxa can undergo extensive hybridization. This random mating can lead to a collapse back into a hybrid swarm (e.g. Katch et al. 2011, Glotzbecker et al. 2016), which can sometimes cause the rarer species to go extinct (see Rhymer and Simberloff 1996). Alternatively, with secondary isolation of a hybrid population, a novel hybrid taxon can be created (e.g. Robertson et al. 2004, Brelsford et al. 2011, Hermansen et al. 2014, Lavretsky et al. 2015, Barrera-Guzmán et al. 2017). A stable hybrid zone (called a tension zone) can also form if ongoing hybridization by parental forms in the hybrid zone counteracts selection against hybrids (e.g. Barton and Hewitt 1985, Moore and Buchanan 1985, Brelsford and Irwin 2009). Within contact zones, hybridization can also allow specific traits to asymmetrically introgress from one species to another (e.g. Stein and Uy 2006, Arnold et al. 2010, Baldassarre et al., 2014, While et al. 2015, Sardell and Uy 2016). Finally, if some pre- or post-zygotic isolation has evolved in allopatry and hybrid fitness is low, the

two taxa in secondary contact can undergo the process of reinforcement (Liou and Price 1994, for a review see Servedio and Noor 2003), where the reduced fitness of hybrids selects for increased prezygotic isolation and divergence of mate recognition signals and/or a narrowing of mate recognition windows (Hudson and Price 2014). Each secondary contact scenario should have specific genomic and phenotypic signatures (see Table 3.1). Therefore, to better understand the speciation process, it is important to examine both the genomic and phenotypic traits of taxa in secondary contact to determine which evolutionary scenario took place, as well as to examine the mechanisms that hinder or allow gene flow during the latter stages of speciation.

Many rapidly speciating lineages and adaptive radiations are restricted to small geographical areas where organisms are isolated from the homogenizing effects of gene flow from other populations and have a variety of open niches to fill (e.g. Darwin's finches in the Galápagos and cichlids in the African Rift lakes; reviewed by Gavrilets and Losos 2009). However, some lineages are able to speciate rapidly while still spreading over a wide geographic area (coined "Great Speciators" by Diamond et al. 1976). This ability to rapidly diversify while covering a wide geographic range is counterintuitive, as high dispersal ability could result in constant gene flow among populations, which should hinder adaptation to local environments and thus halt the evolution of reproductive isolation required for the origin of species (Diamond et al. 1976; see Claramunt et al. 2012, Weeks and Claramunt 2014). This question is known as "the paradox of the great speciator" (Diamond et al. 1976).

Given this paradox, it is essential to investigate how highly dispersing populations remain intact in the face of potential gene flow in secondary contact. One excellent model

to examine speciation in the face of gene flow is the Zosteropidae family of birds. The Zosteropidae family consists of approximately 115 species found in 14 genera, of which 75% of the species fall into the *Zosterops* genus, making it the most species-rich genus of birds (van Balen 2001). *Zosterops* species often have a distinct white eye-ring of small feathers (often known as "White-eyes" or "Silvereyes" colloquially). The Zosteropidae family is one of the most rapidly diversifying vertebrate lineages known, with a radiation of over 100 species in the last two million years (Moyle et al. 2009). This is a per-lineage diversification rate estimated at 1.95-2.63 new species per million years, and is higher or comparable to the per-lineage rate of speciation in other well-known rapidly-diversifying vertebrate lineages, such as Hawaiian Honeycreepers, Darwin's finches, and African cichlids (Moyle et al. 2009). However, compared to these aforementioned lineages, the Zosteropidae clade is the only rapidly speciating lineage found over a very large geographical range, and implies the occurrence of strong colonization ability and yet rapid differentiation in allopatry within this lineage—hence its classification as "a great speciator." Species are found across Africa, Asia, Australia, and through the Indo-Pacific (van Balen 2001). In addition, the increased dispersal ability in this clade is well-documented, as natural range expansions have occurred—for example, Tasmanian populations of *Z. lateralis* naturally colonized New Zealand, Chatham Island, and Norfolk Island sequentially between 1830 and 1904 (North 1904, Falla et al. 1966, Mees 1969, Clegg et al. 2002), and lineages have repeatedly transitioned into new habitats (e.g. Melo et al. 2011, Cox et al. 2014). However, almost nothing is known about the specific reproductive barriers that prevent gene flow and allow lineages to remain distinct between dispersing populations that come into secondary contact in Zosteropidae.

Although several examples exist where multiple White-eye *Zosterops* species or subspecies co-exist and are thought to have come into secondary contact within the last two million years (e.g. Warren et al. 2006, Clegg and Phillimore 2010, Melo et al. 2011, Wickramasinghe et al. 2017), we know of only two previous studies in which gene flow has been examined and hybridization detected between two parapatric subspecies of White-eyes using multiple genomic markers. One is between subspecies of *Z. barbonicus* on Reunion Island (0.434 my diverged; Milá et al. 2010), and the other is in *Z. virens* subspecies in Cape Region of South Africa (0.77 my diverged; Oatley et al. 2012, 2017). To our knowledge, no other study has examined the genomics of reproductive isolation in depth between two overlapping species of *Zosterops* at an age greater than 1 million years apart, or using a genome-wide sequencing method such as RADseq.

On the 15-km wide and 1770-m high Kolombangara Island in the Solomon Archipelago, two overlapping species of White-eye birds have naturally established secondary contact: *Zosterops kulambangrae* and *Z. murphyi*. These two species differ in plumage and song (Diamond 1998, Dutson 2011), are closely-related but not sister taxa (Moyle et al. 2009, Fig. 3.1), and likely arose from two separate colonization events of the Solomon Islands (Mayr and Diamond 2001, Moyle et al. 2009). *Z. kulambangrae* is found in forested habitat from 10 meters up to approximately 1000 meters in elevation, whereas *Z. murphyi*, which is endemic to Kolombangara, is found from approximately 600 meters to the peak of the island at 1770 meters. The ranges of the two species overlap for approximately 400 meters in elevation, from about 600 meters to 1000 meters (Fig. 3.1). This contact zone forms a ring around the island, where the two species can be found together in mixed flocks (Weeks et al. 2017, S. Cowles pers. obs.).

The aim of our study was to examine the genomic consequences of secondary contact between these two overlapping and closely-related *Zosterops* species on Kolombangara. We estimated gene flow across the contact zone, and analyzed differences in vocalizations, plumage, and morphology across individuals of each species. Our ultimate goal was to use a genome-wide sequencing approach to see if these two closely-related *Zosterops* species are reproductively isolated, and to see what factors may be affecting the extent of gene flow within the contact zone. Given the extremely high rate of speciation in the *Zosterops* lineage, we expect that strong reproductive isolation results in limited gene flow between *Z. kulamangrae* and *Z. murphyi*, facilitating their ability to co-exist despite broadly overlapping ranges. In addition, to place our results in a broader, more comparative context, we conducted a literature review to assess the extent of reproductive isolation that occurs in relation to time since divergence between two congeneric avian species in secondary contact (restricted to cases where congeneric taxa are parapatric or have a known/expected hybrid zone).

Methods

Study Site and Fieldwork

We conducted fieldwork from May through June of 2016, and from May through July of 2017 on Kolombangara, a 688-square kilometer round island located in the Western/New Georgia Province of the Solomon Archipelago. The island is an extinct stratovolcano, with three major peaks ranging from 1698 meters to 1770 meters in elevation. Although much of the island's forest has been logged below 400 meters and therefore consists of scrub, secondary regrowth forest, and non-native plantation forest,

400 meters and above is protected and unlogged native hardwood tropical forest. The forest vegetation begins a transition into montane tropical cloud forest at approximately 1000 meters in elevation.

We worked along two transects spanning an elevational range (see Fig. 3.1). One transect started from the village of Iriri (S8° 02.305' E156° 57.484') at 40 meters in elevation, and followed the Iriri Corridor trail to the peak of Mt. Rano up to 1500 meters in elevation (S7° 59.415' E157° 03.454'). The second transect started on the Ringgi Corridor from the road leading up to Imbu Rano Lodge (S8° 02.407' E157° 07.450') at 314 meters in elevation, continued past Imbu Rano Lodge at approximately 360 meters in elevation where the preserved forest begins, and followed along the trail to the peak of Mt. Tepalamenggutu up to 1500 meters in elevation (S7° 58.537' E157° 05.311').

In each year, birds were captured using both six and 12 meter mist-nets along both elevational transects. *Z. kulambangrae* individuals were captured with both passive and target netting using song playback. *Z. murphyi* individuals were caught using passive netting only, given individuals in large flocks (see below) did not respond to playback. Birds were given unique combinations of metal and color leg bands, and blood samples were taken from the brachial vein and stored in lysis buffer (Longmire et al. 1997). Morphometric measurements (beak length, width, and depth, unflattened wing chord, tarsus length, tail length, and mass) were taken from each bird. Three to four feathers were also taken from four body patches on each bird (back, chest, rump, belly) for spectral color analyses. Photos including a mm ruler standard were taken of the left and right white eye ring of each bird to measure eye ring area. All field methods were reviewed and approved by the University of Miami's IACUC. Overall, we captured 76 *Z.*

kulambangrae and 58 *Z. murphyi* along the two transects. During our study period in 2016-2017, most *Z. kulambangrae* individuals that we encountered formed territorial pairs and remained in the same general location. In contrast, most *Z. murphyi* individuals we encountered were transient and would form large flocks that would predictably move along the same routes from high elevations in the morning down through the contact zone throughout the day and back up in the evening. (However, we also observed a few territorial *Z. murphyi* individuals that responded to playback and some transient *Z. kulambangrae* individuals throughout our study—particularly during the later months, suggesting territorial behavior may vary seasonally). Even with these two different behaviors, we caught birds of both species in the same mist nets at the same time within the contact zone and also observed them together in flocks, demonstrating that they were associating with one another within the contact zone. Because of these two different movement behaviors during our study period, we first assessed three genomic groups: low elevation *Z. kulambangrae* (600 meters and below), contact zone *Z. kulambangrae* (600-1000 meters), and all *Z. murphyi* (600 meters and above). Of the 76 *Z. kulambangrae*, 20 birds were captured within the contact zone from 600-1000 meters in elevation, leading to 56 low-elevation *Z. kulambangrae* (below 600 meters), 20 contact zone *Z. kulambangrae* (600-1000 meters), and 58 *Z. murphyi* (600-1500 meters).

GPS points were taken for each captured bird. To assess climate in each species' habitat, we downloaded data from the Worldclim2 database (Fick and Hijmans 2017) for each GPS point where birds were captured using the 'raster' package in R (Hijmans and van Etten 2014). We used a spatial resolution of 0.5 minutes. We then calculated the temperature and precipitation ranges and means for locations where we caught

individuals of each species. For *Z. kulambangrae*, the average temperature ranged from 21.4-27° C (mean ± SE = 24.7° ± 1.6° C) and the annual precipitation ranged from 3637-3982 mm (mean ± SE =3779.6 ± 10.0 mm). For *Z. murphyi*, the average temperature ranged from 20.0-24.4 ° C (mean ± SE = 20.8° ± 1.1° C) and the annual precipitation ranged from 3787-4081 mm (mean ± SE = 4027.9 ±7.8 mm).

Mitochondrial and Genomic Sequencing

DNA was extracted from blood samples using Qiagen's DNeasy Blood and Tissue Extraction Kits (Qiagen, Hilden, Germany) following the standard extraction protocol for nucleated blood. For mitochondrial DNA sequencing, we followed a standard PCR amplification protocol using the H1064 and L5125 primers developed for sequencing the mitochondrial gene NADH dehydrogenase subunit 2 (ND2) in birds (Sorenson et al. 1999), with an annealing temperature of 54° C. PCR-amplified DNA was cleaned using the manufacturer's ExoSAP-IT protocol (USB corporation, Cleveland, OH, USA), prepared for sequencing using the BigDyeTerminator Version 3.1 Cycle Sequencing Kit (Applied Biosystems, Waltham, MA, USA), and purified with Sephadex (Sigma-Aldrich, St. Louis, MO, USA). Samples were loaded onto 96-well plates and were sequenced using Sanger Sequencing at the Hussman Institute for Human Genomics at the University of Miami's Medical Campus (Miami, FL, USA) and at the Molecular Core Facility in the Department of Biology (Coral Gables, FL, USA). Mitochondrial sequences were visually inspected and subsequently aligned using Sequencher version 4.8 (Gene Codes, Ann Arbor, MI, USA). We then created median joining haplotype networks using the software program PopART version 1.7 (Leigh and Bryant 2015).

For genomic sequencing, DNA quality was checked by running 100ng of each sample on a 1% agarose gel with electrophoresis to confirm intact, non-fragmented DNA. We then digested 300ng of 10% of the samples with the manufacturer's EcoRI digestion protocol (Promega, Fitchburg, WI, USA) to confirm enzymatic activity. DNA samples were loaded onto 96-well plates and sent to the University of Wisconsin-Madison's Bioinformatics Resource Center for genotyping-by-sequencing using the restriction enzyme ApeKI. SNPs were then called by the University of Wisconsin-Madison's Bioinformatics Resource Center using the established TASSEL pipeline and parameters (Glaubitz et al. 2014) and subsequently mapped to the Zebrafinch (*Taeniopygia guttata*) ENSEMBL genome assembly (taeGut3.2.4.87). Briefly, this TASSEL pipeline calls SNPs with a minimum minor allele frequency of 0.01, a minimum minor allele count of 10, a mismatch rate of below 0.1, and a minimum site coverage of 0.2. The filtered dataset ended up being 79,742 mapped SNPs. However, for our subsequent genomic analyses, we used a reduced dataset of 23,752 SNPs that had a minor allele frequency of 0.05 and were found in at least 70 of the 134 sequenced individuals of both species. We then created a vcf file with these 23,752 SNPs in TASSEL version 1.5 (Bradbury et al. 2007) for use with later genomic analyses.

Measures and Visualization of Genomic Gene Flow

To infer gene flow between the two species, we used several genomics programs designed to work with RADseq SNP data. First, we used the Bayesian-based program fastStructure (Raj et al. 2014) to examine population structuring using our filtered GBS dataset of 23,752 SNPs. To use this program, we converted our vcf file to plink bed

format using the program PLINK version 1.07 (Purcell et al. 2007) and input the bed, bim, and fam files to fastStructure. In separate model runs, we changed the number of assigned populations (K) from 2 to 6. We used a convergence criterion of $10e^{-8}$, 5 cross-validation sets, and the simple prior (flat-beta prior). From our output runs, we used the 'chooseK.py' script available in fastStructure, which uses a marginal likelihood estimate on multiple model runs with different K values to choose the best-fit model to our dataset.

Second, we calculated Weir and Cockerham's F_{ST} values for each SNP using the 'hierfstat' package in R (Goudet 2005, Goudet and Jombart 2015). For this analysis, we first converted a diploid hapmap file of our dataset from TASSEL version 1.5 (Bradbury et al. 2007) into a formatted text file for R by converting bases into numerical digits and recoding missing values as 'NA.' We used species identity (i.e. either *Z. kulambangrae* or *Z. murphyi*) as the level for calculating F_{ST} values for each SNP locus. Next, we averaged SNP F_{ST} values across 100kb windows for each chromosome using a custom script written in R v3.3.2 (R Development Core Team 2016). To visualize F_{ST} values across the genome, we plotted these 100kb-window F_{ST} averages on a Manhattan plot using the R package 'qqman' (Turner 2014).

Finally, we conducted a principal component analysis (PCA) of our genomic data in the R package 'adegenet' v.2.0.2 (Jombart and Ahmed 2011). We first converted our vcf file into a genind object using the R package 'vcfR' (Knaus and Grunwald 2017), and then ran a PCA, with missing values substituted by the mean (NA.method = mean). We then plotted the eigenvalues for PC1 vs PC2 for each individual in R v3.3.2 (R Development Core Team 2016).

Song and Call Analyses

Song and calls of both *Zosterop*s species were recorded in 2016-2017 along both transects with two separate types of recorder: a Marantz (Mahwah, NJ, USA) PMD-670 solid state recorder and an Olympus (South Hackensack, NJ, USA) LS-12 Linear PCM Digital voice recorder, each attached to the same Sennheiser (Old Lyme, CT, USA) MKE 660 shotgun microphone. Songs were recorded opportunistically when birds were singing or calling in close proximity in the morning from dawn (approximately 6:30am) to 4pm. Overall, we were able to record 7-24 clear songs with minimal background noise from 8 *Z. kulambangrae* and 3 *Z. murphyi* individuals. We also recorded 10-20 clear calls from 4 *Z. kulambangrae* and 4 *Z. murphyi* individuals or groups of individuals. Given that birds usually call in pairs or when in large groups (particularly in the case of *Z. murphyi*), it is often difficult to record calls of single individuals. Given our inability to isolate individuals, we treated each pair/group call recording as a single sample even if multiple individuals were calling together. Recordings of non-banded individuals were done in succession to make sure we recorded the same individual or individuals for the entire recording. We marked each recording spot with a GPS point to reduce the possibility of returning to the same bird or group of birds. We also recorded songs from individuals more than 600 meters apart and calls from pairs/groups more than 1000 meters apart to make sure we recorded different individuals each time.

We used the program Raven Pro Version 1.5 (Bioacoustics Research Program 2014) to make spectrograms of each clear song and call recording. From each song, we took measurements of the number of syllables per song, min frequency, max frequency, peak frequency (the frequency with the largest amplitude), 90% bandwidth (difference

between 95% and 5% frequencies), and song duration. We averaged these measurements per individual for comparison within and across species and conducted a PCA on the correlation matrix of the song variables in R v3.3.2 (R Development Core Team 2016). For each call, we took measurements of low frequency, high frequency, peak frequency, 90% bandwidth, and call duration. We averaged these measurements per individual or group sample for comparison and conducted a PCA on the correlation matrix of the call variables in R v3.3.2 (R Development Core Team 2016). Subsequently, we used Mann-Whitney U tests (due to uneven and low sample sizes for both calls and song) to determine if the song and call principal components were significantly different from one another.

Plumage Color and Eye-Ring Analyses

We obtained a set of three-four feathers from four distinct body patches (back, chest, rump, belly) on all captured birds of both species. Feathers in each set were stacked on top of each other and mounted on Strathmore (West Springfield, MA, USA) 400 Series Artagain drawing paper in Coal Black. To obtain an objective measure of color, we used an Ocean Optics (Largo, FL, USA) USB2000 Spectrophotometer connected to a pulsed Xenon PX-2 light source (Ocean Optics, Largo, FL, USA). We took a reflectance reading from 300-700nm from each feather sample using the OOIbase32 software (Ocean Optics, Largo, FL, USA). Reflectance readings were standardized at a distance of 1 cm from the feather samples using a black anodized aluminum probe attachment angled at 45 degrees (to reduced specular glare, see Endler 1990).

To compare color across body patches and individuals, we used the vismodel function in the R package 'pavo' (Maia et al. 2013). This function calculates the quantum catch of each spectral curve for the four chromatic photoreceptors in an avian visual system based on the models of Vorobyev and Osorio (1998). We used the chromatic parameters specified for blue tits (used previously for *Zosterops* in Cornuault et al. 2015). We then calculated just-noticeable difference (JND) measurements for each of the four body patches for individuals in avian color space using the 'coldist' function and standard parameters. A single just-noticeable difference unit is equal to the threshold that an individual bird can perceive a difference in color at least 50% of the time (Weber 1834). Pairwise comparisons of chromatic distance JNDs were calculated for *Z. kulambangrae* and *Z. murphyi* individuals both within and across the species level and visualized in R v3.3.2 (R Development Core Team 2016). Additionally, we conducted a PCA of the color wavelengths (standardized for brightness) in R v3.3.2 (R Development Core Team 2016) and used t-tests to determine if the principal components were significantly different from one another.

We also took digital images of the left and right white eye-rings of each captured individual next to a mm ruler standard in a horizontal plane. We used the program ImageJ 1.32 (Schneider et al. 2012) to calculate the area in square millimeters of the left and right eye rings after converting the image to an 8-bit greyscale mask. We averaged the right and left eye-ring area measurements per individual so that we had only one area measurement per individual. We then plotted the averages of the eye-ring sizes using the boxplot function and conducted a t-test in R v3.3.2 (R Development Core Team 2016) to see if the average eye-ring size is different between the two species.

Literature review to examine avian gene flow in secondary contact

Our aim was to compare our calculated level of reproductive isolation among the two Kolombangara *Zosterops* species with reproductive isolation levels in other pairs of congeneric avian species or subspecies known to have established secondary contact. We conducted a literature review to find examples of known presence or absence of avian gene flow and hybridization in secondary contact (defined as congeneric taxa with a parapatric distribution or a known/expected hybrid zone). We also used the comprehensive avian hybrid zone list provided in Chapter 15 of Price (2008). Overall, we found a total of 88 examples where either gene flow or a phenotypic measurement of hybridization was measured, and the time of divergence was estimated between two congeneric avian species or subspecies in secondary contact. If only an estimate of percent mitochondrial divergence was given and not a specific time, we used the standard estimate of 2% divergence rate per million years for mtDNA in birds (similar to Price 2008). We used a measure of hybrids (both phenotypic and/or genetic) detected per the number of individuals sampled, and subtracted this number from 1, giving us reproductive isolation indices that ranged from 0 (no reproductive isolation detected at all; all individuals detected at the center of the contact zone were phenotypically or genomically hybrids) to 1 (zero hybridization and complete reproductive isolation detected in the contact zone). When more than one estimate was given for hybridization based on different methods or studies, we used the more conservative estimate. We plotted all examples in R v3.3.2 (R Development Core Team 2016). To further investigate the relationship between time and reproductive isolation, we conducted logistic regression models in R using the glm function (family = binomial) (R

Development Core Team 2016). We fit a logistic regression model to the entire dataset (all 88 examples), both with and without a forced intercept of 0,0 (i.e., the assumption that 0 time is associated with 0% reproductive isolation). We also fit logistic regression models to the subcategories based on the method of assessing gene flow in hybrids (both genetic and phenotypic methods (50 examples), phenotypic methods only (7 examples), and genetic methods only (31 examples)).

All the examples we found were from systems known to have parapatric distributions (and therefore defined contact/hybrid zones). In addition, we included examples in which distinct mtDNA haplotypes have been discovered within the range of a single species (i.e., they fit a fusion scenario). We acknowledge that there is likely a publication and study bias for examples of two congeneric species that are suspected of hybridizing, as well as a paucity of studies in which overlapping or sympatric congeners exist but do not hybridize (given these situations are not usually tested for hybridization and gene flow). However, despite this limitation, we believe that the published literature can provide valuable information about the time it takes to evolve reproductive isolation in birds.

Results

Mitochondrial and Genomic Gene Flow Between Species

A total of 744 bp of the mitochondrial gene *ND2* were successfully sequenced for 75 *Z. kulambangrae* and 56 *Z. murphyi* individuals. Overall, the median-spanning mitochondrial haplotype network showed distinct species separation, with over 30 sequence changes (36 between the two closest individuals) between the two species (Fig.

3.2a). In addition, all *Z. kulambangrae* contact zone birds clustered within all non-contact zone *Z. kulambangrae* (Fig. 3.2a, purple and blue, respectively). Using a conservative ND2 clock estimate of 0.02-0.022 changes per base pair per million years (Arbogast et al. 2006) and a total sequence divergence of 4-5%, we estimate that the most recent common ancestor between these two species was approximately 2.2-2.42 million years ago.

Using the program fastStructure, all individuals were identified as either 100% *Z. kulambangrae* or 100% *Z. murphyi* (Fig. 3.2b), with all fastStructure q-values (often used as a hybridization index) < 0.001 for all *Z. kulambangrae* individuals and > 0.999 for all *Z murphyi* individuals. K=2 was the best-fit model using the fastSTRUCTURE script 'choose.py,' and changing the number of iterations or convergence parameters of the original model did not affect the population structuring results. We also calculated Weir and Cockerham's F_{ST} values for our dataset of 23,752 SNPs, and found SNPs with high F_{ST} values scattered throughout the genome, consistent with no gene flow (Fig. 3.2c). A total of 200 SNPs scattered throughout the genome had an F_{ST} value equal to one, and the average F_{ST} value per SNP for all 23,752 SNPs in the dataset was 0.154 across the entire genome (Fig. 3.2c). A principal components analysis in the R package 'adegenet' showed that the first principal component explained 5.66% of the genomic variation and clearly separated the two *Zosterops* species (Fig. 3.2d). All contact-zone and non-contact zone *Z. kulambangrae* individuals clustered together in the genomic PCA (Fig. 3.2d, purple and blue, respectively). For all subsequent analyses, we only used two species-specific groupings (i.e., *Z. kulambangrae* and *Z. murphyi*) as distinct groups, given *Z. kulambangrae* contact zone birds are not genetically distinct from *Z. kulambangrae* non-contact zone birds.

Morphometric Measurements

We compared seven morphometric measurements between the two *Zosterops* species. *Z. murphyi* was significantly larger than *Z. kulambangrae* in six of the seven measurements (all but beak width, see Table 3.2), which suggests that *Z. murphyi* individuals are generally larger overall than *Z. kulambangrae* individuals.

Species-specific Song and Call Characteristics

We analyzed between 7-24 songs per individual for eight *Z. kulambangrae* and three *Z. murphyi* individuals. Overall, we found that a principal components analysis was able to distinguish the two species' songs in sound space (Fig. 3.3). PC1 explained 71.8% of the variation, and was negatively associated with min frequency [-0.44] and positively associated with all other song variables measured (max frequency [0.46], peak frequency [0.45], 90% bandwidth [0.42], duration [0.2], and number of syllables [0.41]). PC1 was significantly different for the two species' songs (Mann-Whitney U test: $U= 0$, $p = 0.012$). PC2 explained 18.1% of the variation and was negatively associated with duration [-0.86], min frequency [-0.07], and number of syllables [-0.35], and positively associated with max frequency [0.21], peak frequency [0.19], and 90% bandwidth [0.24]. PC2 was not significantly different for songs between the two species (Mann-Whitney U test: $U=8$, $p = 0.5$).

Similarly, we found that the two species have distinctive calls in sound space (Fig. 3.3). We analyzed 10-12 calls from four individuals/groups of each species. Using a principal components analysis, PC1 explained 53.8% of the variation, and was positively associated with all measured variables (min frequency [0.54], max frequency [0.52], peak

frequency [0.41], 90% bandwidth [0.43], and duration [0.31]). PC1 was significantly different for the two species' calls (Mann-Whitney U test: $U = 0$, $p = 0.03$). PC2 explained 27.9% of the variation and was positively associated with max frequency [0.41] and 90% bandwidth [0.59], and negatively associated with duration [-0.34], min frequency [-0.25], and peak frequency [-0.55]. PC2 was not significantly different for calls between the two species (Mann-Whitney U test: $U = 5$, $p = 0.49$).

Plumage Color and Eye-Ring Size Comparison

To compare overall plumage color between the two species, we compared feathers from four separate body patches: back, belly, chest, and rump. Using a spectrometer to take reflectance readings from 300-700nm, we found that overall color spectral curve shape was qualitatively similar between the two species for chest and back (Fig. 3.4; belly and rump spectral curves are similar and are not shown). Using a PCA on the wavelengths for these two body patches, we found that PC1 and PC2 were significantly different between *Z. kulambangrae* and *Z. murphyi* for back (PC1 t-test: $t = 6.73$, $p<0.001$; PC2 t-test: $t = 6.51$, $p<0.001$) and chest (PC1 t-test: $t = 5.05$, $p <0.001$; PC2 t-test: $t = 2.79$, $p = 0.006$; Fig. 3.4), although there was much overlap between species in PC1 and PC2 scores for both patches (Fig. 3.4). In addition, to obtain a biologically realistic measure of visible color difference, we calculated the color JNDs of all four plumage patches both and within and across species using an avian visual model in the R package 'pavo'. We found that both within and across species color variation was similar in magnitude and around the value of one JND (range: 0.90-1.22) for all four body patches (Fig. 3.4).

In contrast, we found that the white eye ring size was different between the two species (Fig. 3.4). Mean $\pm$ SE white-eye ring size was 9.93 $\pm$ 0.30 sq mm for *Z. kulambangrae*, and 44.76 $\pm$ 1.01 sq mm for *Z. murphyi*. These measurements were significantly different between species (t-test: $t = -32.82$, $p < 0.001$).

Overview of avian gene flow in secondary contact

We conducted a literature review to obtain a comprehensive overview of the time it takes to establish complete reproductive isolation (i.e., absence of detectable hybrids) in secondary contact between congeneric avian species or subspecies with parapatric distributions or an established hybrid zone. Using detected phenotypic and genomic hybrids as a level of hybridization in secondary contact, we found that the total level of gene flow across congeneric pairs of bird species in secondary contact ranged from 0-100%, and spanned a timeframe of 0.12-6.95 million years ago (Fig. 3.5, Table 3.3). Our study case of the two Kolombangara *Zosterops* species was the youngest known case to date of zero gene flow and complete reproductive isolation detected between two congeneric species using genomic data, in which the two species have overlapping breeding ranges (Fig. 3.5). The only other younger known case of complete reproductive isolation upon secondary contact we found is in two subspecies of Willets (*Tringa semipalmata semipalmata* and *T. s. inornata*) that are only 700,000 years apart; however the two species do not overlap in their breeding ranges (Fig. 3.5, Table 3.3, for details see Oswald et al. 2016). The logistic regression models based on reproductive isolation indices calculated from all 88 cases, or subcategories including genetic data (both genetic and phenotypic methods—50 examples, and genetic only—31 examples) had time as a

significant or near-significant predictor of reproductive isolation, but only in the forced-intercept models (see Table 3.4).

Discussion

Absence of gene flow suggests strong reproductive barriers

Despite recent divergence and extensive range overlap, we detected no gene flow between the two species of *Zosterops* White-eyes on the island of Kolombangara along our two replicate transects. This absence of gene flow between these two species is noteworthy, given that it takes approximately five million years for hybrid infertility, and on average twice as long for hybrid inviability to evolve in passerines (Price and Bouvier 2002, Price 2008). In addition, our estimate of a divergence time of 2.2-2.42 mya is extremely conservative for species in this lineage, given the young age of several of the neighboring islands in the New Georgia Province (estimated as young as 560,000 years old)—each with a different closely-related congeneric *Zosterops* species (Moyle et al. 2009) with distinct plumage and song (Dutson 2011, Diamond 1998). Given that hybridization is common in birds (Grant and Grant 1992), detectable in several other congeneric avian species in secondary contact with similar or greater times of divergence (see Fig. 3.5, Table 3.1), and in younger white-eye secondary contact zones less than 0.8 million years old (e.g. Milá et al. 2010, Oatley et al. 2012, 2017), our failure to detect gene flow between these two *Zosterops* species suggests that the presence of strong reproductive barriers existed or rapidly evolved upon secondary contact (i.e., Scenario 1 or 2 from Table 3.1).

Even though we cannot conclusively determine the mechanisms preventing gene flow in these two species, strong barriers preventing hybridization and reducing hybrid fitness are likely critical as both premating and postmating barriers together are likely needed to drive complete reproductive isolation in birds (Uy et al. 2018). Given that birds use song and plumage as traits in species recognition (reviewed in Uy et al. 2018), and changes in these traits have been shown to correlate with reproductive isolation and speciation rates in birds (e.g. Searcy 1990, Barraclough et al. 1995, Grant and Grant 1996, Irwin et al. 2001, Uy et al. 2009, Hugall and Stuart-Fox 2012, Seddon et al. 2013), the notable differences in visual and acoustic signals between these two *Zosterops* species suggest plumage and song may be important for pre-mating isolating.

First, given that the calls and songs are distinct between the two Kolombangara *Zosterops* species, vocalizations are likely important premating barriers. Vocalizations in birds are likely used as a long-range communication signal between individuals (Uy and Safran 2013), which may be particularly relevant to this species pair given they are found in dense forested habitat. In other *Zosterops* White-eyes, songs and calls show acoustic characteristics that match their habitats (Potvin et al. 2011, Potvin and Parris 2012), are known to be variable at small scales (Habel et al. 2015), and vary across closely-related species (Diamond 1998). Also, character displacement has been found in the calls of two parapatric East African *Zosterops* White-eyes (Husemann et al. 2014, Habel et al. 2015). These studies provide evidence for the importance of species-specific and habitat-specific vocal communication in *Zosterops*. Although our sample sizes are limited, we found that the two Kolombangara *Zosterops* species do have distinct songs and calls that do not

overlap in sound space, suggesting that songs and calls could be important in conspecific recognition.

In contrast, body plumage coloration might be less important as a species-recognition signal between these two species given that only small color differences exist between *Z. murphyi* and *Z. kulambangrae*, and the level of across-species chromatic variation is similar to within-species chromatic variation using an avian visual model (Fig. 3.4). Other studies on *Zosterops* species have found that greater plumage variation exists than would be expected by chance in closely-related lineages in secondary contact (e.g. Cornuault et al. 2015, Melo et al. 2011) or that plumage can rapidly evolve in short times scales (Milá et al. 2010, Melo et al. 2011), which suggests plumage may play a more important role in other *Zosterops* species pairs in secondary contact. However, eye ring-size was extremely different between the two species and could also be used as a species-recognition signal—perhaps as a short-range signal (Uy and Safran 2013). The existence of two sympatric *Zosterops* species on several islands in the Gulf of Guinea, one of the "typical" White-eye form with an eye-ring and one "aberrant" form without an eye ring (in addition to other differences in body coloration, see Melo et al. 2011) lends support to this idea. But, as far as we know, no study has examined the function of eye ring plumage in *Zosterop*s species or in other birds.

Hudson and Price (2014) also suggest that factors other than the sexual signals themselves, such as the narrowing of the female recognition window or the rarity of mates could affect the level of hybridization upon secondary contact. They found that song and plumage divergence were not strong predictors of premating isolation in their review of 17 well-studied avian hybrid zones, which points to the possibility that

reinforcement of the female recognition window may play an important role in establishing reproductive isolation. As far as we know, no study has examined the possibility of a narrowed female response window in the *Zosterops* genus.

It is also possible that the two *Zosterops* species on Kolombangara do not overlap in breeding range. If *Z. kulambangrae* only breeds below 600 meters, and *Z. murphyi* only breeds above 1000 meters (as noted in Weeks et al. 2017), this spatial separation of breeding range could effectively prevent any hybridization. The transition starting at 1000 meters from tropical montane forest to cloud forest on Kolombangara is consistent with this hypothesis. However, we did note males of both species acted territorial (remained in the area and sang repeatedly) in the upper region of the contact zone (around 1000m), suggesting the possibility for at least some territorial overlap during the breeding season. In addition, the timing and duration of the breeding season for both species is completely unknown. If the breeding season for each species does not overlap, this could also prevent hybridization. Local habitat and climate differences can drive asynchronous breeding times even within close populations of the same bird species (e.g. Moore et al. 2005, Caro et al. 2009). The temperature and precipitation differences in the lowland tropical montane forest versus higher elevation cloud forest on Kolombangara could lead to different optimal breeding times for each species.

Post-mating barriers could also play an important role in preventing gene flow between these two Kolombangara *Zosterops* species through reduced hybrid fitness. The climatic variation between the species' habitats should favor unique morphological and physiological adaptations, and it is possible that hybrids would suffer from being poorly adapted to either environment. Sandvig et al. (2017) found that survival of *Zosteropos*

lateralis on a small island near Australia was indeed correlated to climatic (rainfall) variables. Additionally, Olson et al. (2010) found that hybrid chickadees (hybrids of *Poecile atricapillus* and *P. carolinensis*) had reduced metabolic efficiency compared to their parental species, and suggested that genetic incompatibilities between mismatched mtDNA and nuclear DNA could have severe physiological costs in hybrid birds. However, in species with strong selection against hybrids, it could be expected that hybrids would be detected but that backcrossed and introgressed individuals would not be found (as in Steeves et al. 2010, Hansson et al. 2012). Since we did not find any hybrids, our results suggest that premating isolation may be more important in reproductive isolation between these two species.

It has also been suggested that White-eye species have elevated rates of genomic evolution (Cornetti et al. 2015), which is consistent with the idea that rates of genomic evolution are increased in rapidly-diversifying avian lineages (e.g. Lanfear et al. 2010, Nabholz et al. 2011). Given that these two species belong to the rapidly-diversifying *Zosterops* lineage, it is possible these two species could have increased genomic differences and therefore increased rates of genetic incompatibilities compared to other bird lineages even though they have only been separated for approximately two million years. However, we did estimate the time of divergence for these two species using changes in the mitochondrial genome, which should have controlled for any increased genomic evolutionary rate in our time estimation (i.e., our estimation of 2 million years is conservative). But again, given that we detected zero hybrids or backcrossed individuals in our sampling, and that on average it takes 10 million years to evolve complete hybrid

inviability in birds (Price and Bouvier 2002), our results suggest that premating isolation between these two species is extremely strong.

Review Implications

In our literature review examining the time it takes to evolve reproductive isolation in congeneric bird species or subspecies with a parapatric distribution or known hybrid zone, we found cases spanning from taxa as recently diverged as 9,000 years ago that show high levels of reproductive isolation (Carrion Crow and Hooded Crow, RI index: 0.88) to taxa close to 6-7 million years apart that are still hybridizing (e.g. Sooty and Cardinal Myzomela, RI index: 0.78; Barred Owl and Spotted Owl, RI: 0.96). These ranges show that in birds, there is likely a wide variety of premating, postmating, and ecological factors important in determining the level of reproductive isolation between two congeneric species that come into secondary contact. In addition, time was a significant or near-significant predictor of reproductive isolation only in the logistic regression models that included genetic data and with a forced intercept at 0. These results suggest that estimates of hybridization/reproductive isolation may be stronger using genetic data, or a combination of genetic and phenotypic data, rather than just phenotypic data alone.

We also show that 1) even at low divergence times of less than 1-2 million years, two congeneric species or subspecies can fall along a wide range of levels of reproductive isolation (i.e., from 0 to 1), and 2) most examples of congeneric species pairs older than 3 million years apart showed higher levels of reproductive isolation. These results again suggest that a multitude of factors likely affect hybridization in taxa that are recently

diverged, but that time since separation increases the extent of reproductive isolation between them. Particularly in birds, a divergence time of at least 2-3 million years between closely-related species may be necessary for establishing overlapping ranges. Hudson and Price (2014) show that on average, allopatric pairs of sister taxa in songbirds are younger in age (1.8 my) compared to sympatric sister taxa (2.3 my). This study also provides evidence that several million years of divergence are necessary for closely-related taxa to accumulate enough reproductive isolation to establish sympatry.

Final Summary

Overall, the two closely-related *Zosterops* species on the island of Kolombangara show no detectable hybridization or genetic introgression within their 400-meter elevational contact zone around the island even though they are only approximately 2.2 million years apart. This outcome is unusual, given that many other bird species in secondary contact are able to form viable hybrids at the same or increased times since divergence (Fig. 3.5, Table 3.3), and also suggests that premating barriers between these two *Zosterops* species are especially strong. Barriers preventing gene flow could include plumage-related traits (specifically eye-ring size), vocalizations (both songs and calls), physiological traits, a lack of overlap in breeding range or timing, and genetic incompatibilities. In addition, our results may further help resolve the "great speciator paradox"—lineages that are widespread and vagile can have high rates of speciation if they can evolve complete reproductive isolation more quickly than other bird lineages. This ability is what may allow recently diverged species to remain distinct upon secondary contact, ultimately allowing these lineages to rapidly speciate yet still spread

over a huge geographic area. Further work examining the specific barriers that hinder gene flow in secondary contact in the rapidly-speciating *Zosterops* genus of birds could improve our understanding of the role of premating and post-mating isolating barriers upon secondary contact in rapidly-speciating yet widespread lineages.

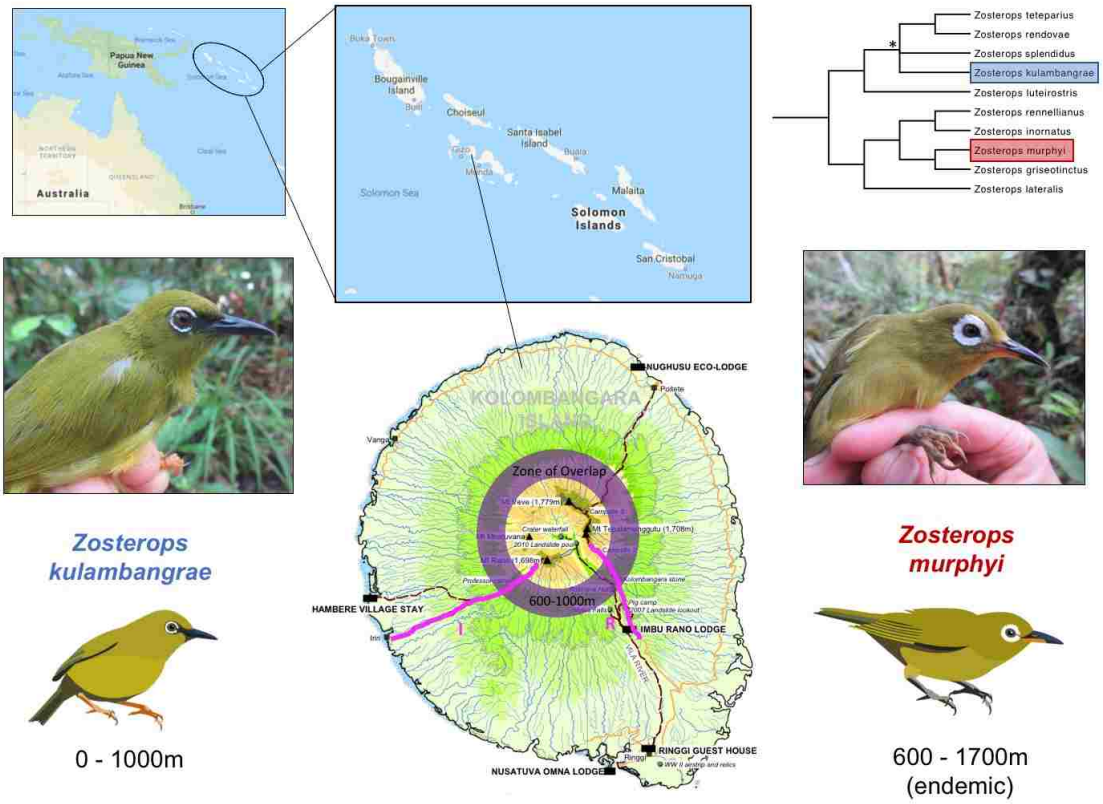

Figure 3.1: Distributional Range. The two *Zosterops* White-eye species, *Z. kulambangrae* and *Z. murphyi*, are found on the island of Kolombangara in the Western Province of the Solomon Islands. The two species have a zone of overlap from approximately 600-1000 meters in elevation (shown in purple), which forms a contact zone ring around the island. However, the two species are not sister taxa. Phylogeny is modified from Moyle et al. (2009), asterisk denotes a calibrated time point of 560,000 years ago. Our two elevational transects are outlined in pink (I = Iriri transect, R = Ringgi transect). Kolombangara map provided courtesy of KIBCA.

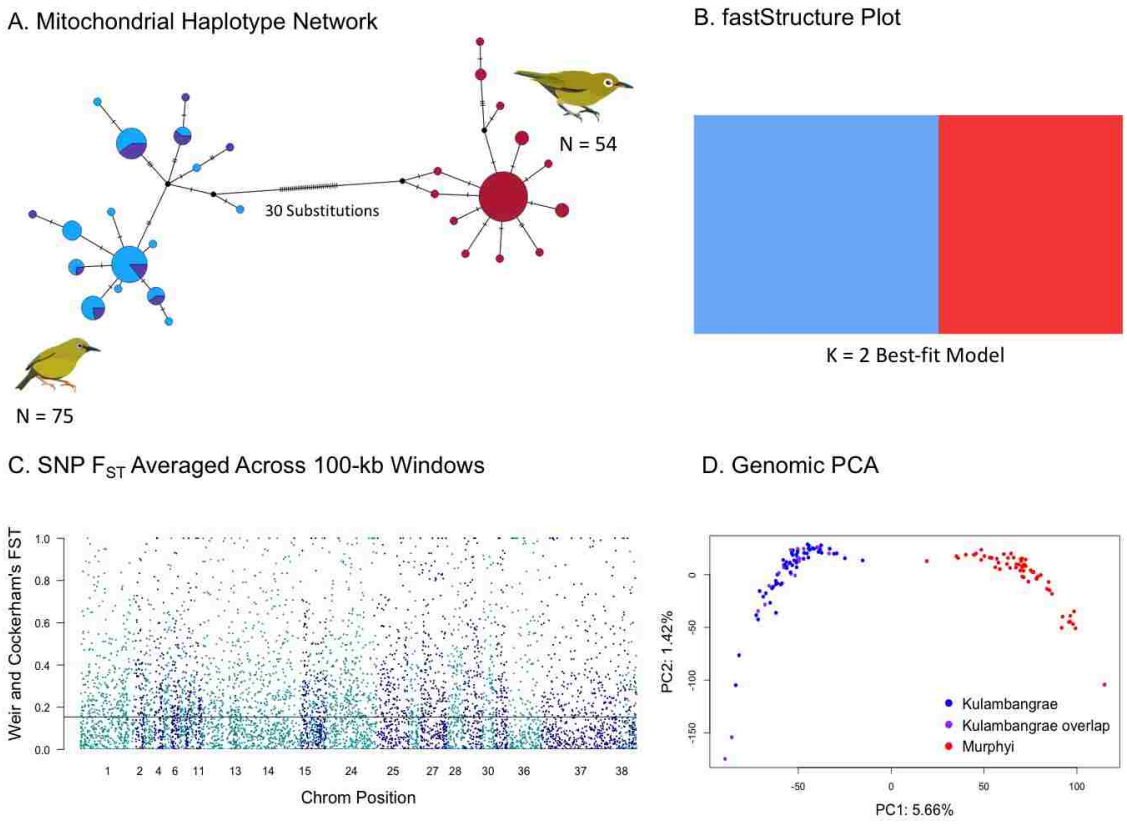

Figure 3.2: Genomic analyses suggest the absence of gene flow between the two *Zosterops* species. A) A mitochondrial haplotype network based on 744bp of the *ND2* mitochondrial gene has over 30 different substitutions between the two different species (blue = *Z. kulambangrae*, purple= contact zone *Z. kulambangrae*, red = *Z. murphyi*). B) The best-fit K=2 fastStructure model shows two distinct group assignments, corresponding with 100% true species identity (blue = *Z. kulambangrae*, red = *Z. murphyi*). C) Weir and Cockerham's F_{ST} values for 23,752 SNPs were averaged in 100kb-windows across the genome. The average F_{ST} value for all SNPs is 0.154, denoted by the black line. D) A principal components analysis on the genomic dataset shows two distinct species clusters that separate on the first principal component axis (PC1: x-axis), which explains 5.66% of the variation in the dataset (blue = *Z. kulambangrae*, purple= contact zone *Z. kulambangrae*, red = *Z .murphyi*). All four panels are consistent with no gene flow between the two species.

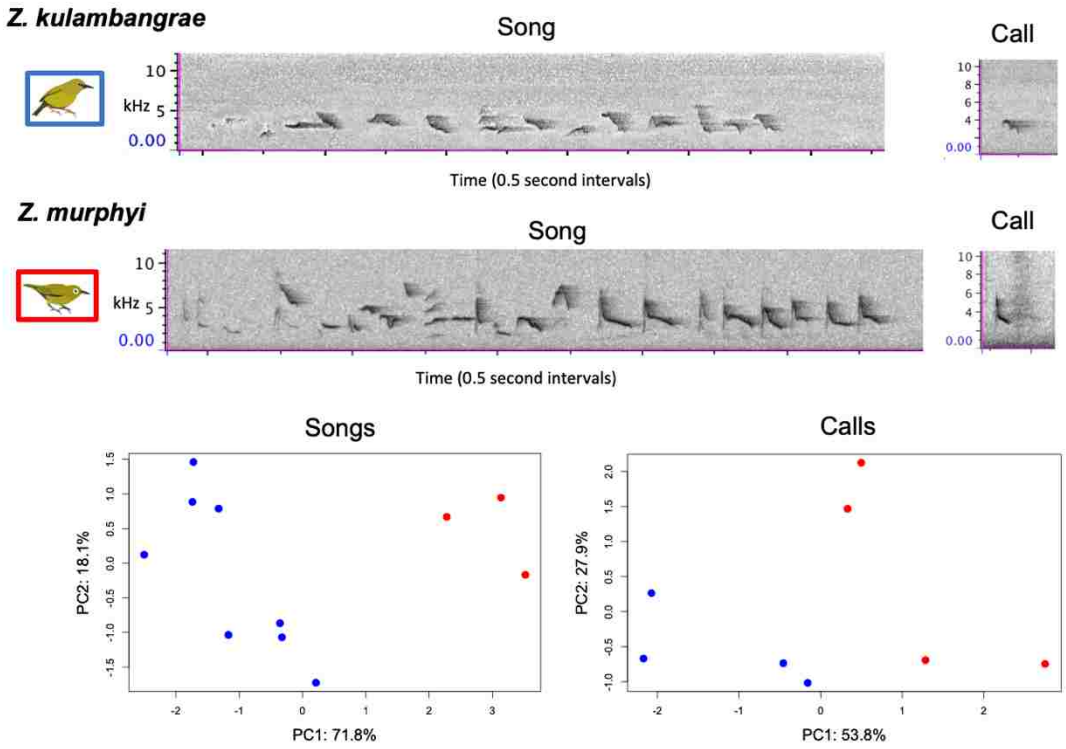

Figure 3.3: Songs and calls are distinct between the two species. Spectrograms of representative songs and calls (upper panels) for each species are shown. Lower panels denote principal components analyses based on song characteristics (number of syllables per song, min frequency, max frequency, peak frequency, 90% bandwidth, and song duration) and call characteristics (min frequency, max frequency, peak frequency, 90% bandwidth, and call duration) for each species. For both song and calls, PC1 is significantly different between the two species, whereas PC2 is not (see text for details).

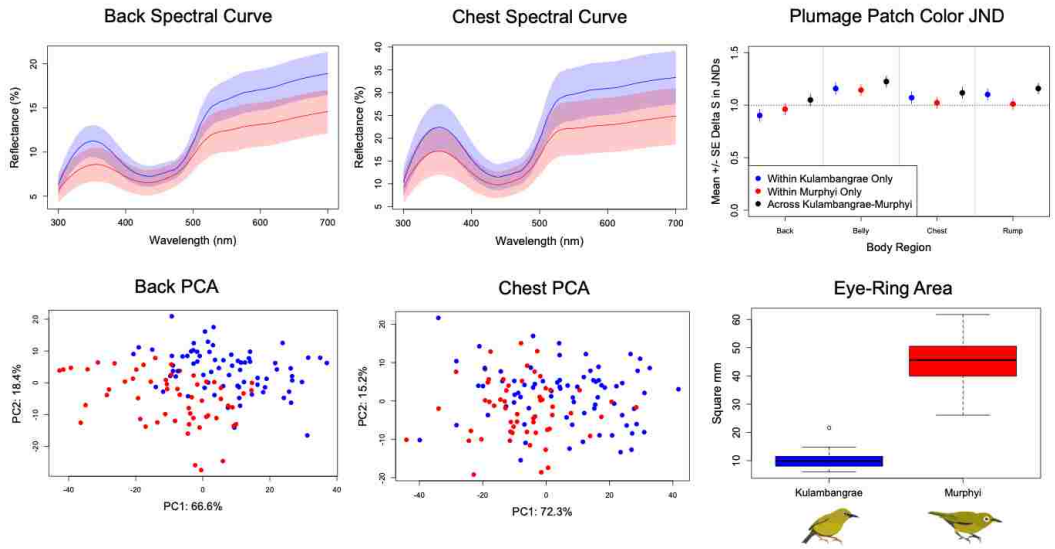

Figure 3.4: Plumage color is similar between the two *Zosterops* species but eye ring size is distinct. We took objective color measurements of feathers from four body patches (back, belly, chest, rump) collected from each captured bird. Shown are the 300-700nm spectral curves (upper left and center panels) and principal components analyses (lower left and center panels) for the back and chest patches, respectively. (Belly and rump are similar and not shown). Blue denotes *Z. kulambangrae*, and red denotes *Z. murphyi*. Comparisons of pairwise plumage patch just-noticeable differences (JNDs) are shown for within each species and comparatively across both species (upper right panel); error bars denote standard error and dotted line denotes the detectable threshold of JND=1. Boxplots of average eye-ring area for each species are also shown (lower-right panel); whiskers denote 95% confidence intervals. Both PC1 and PC2 for chest and back are significantly different between the two species (see text for details), but the JND measurements from the avian visual model suggest the plumage patches appear quite similar.

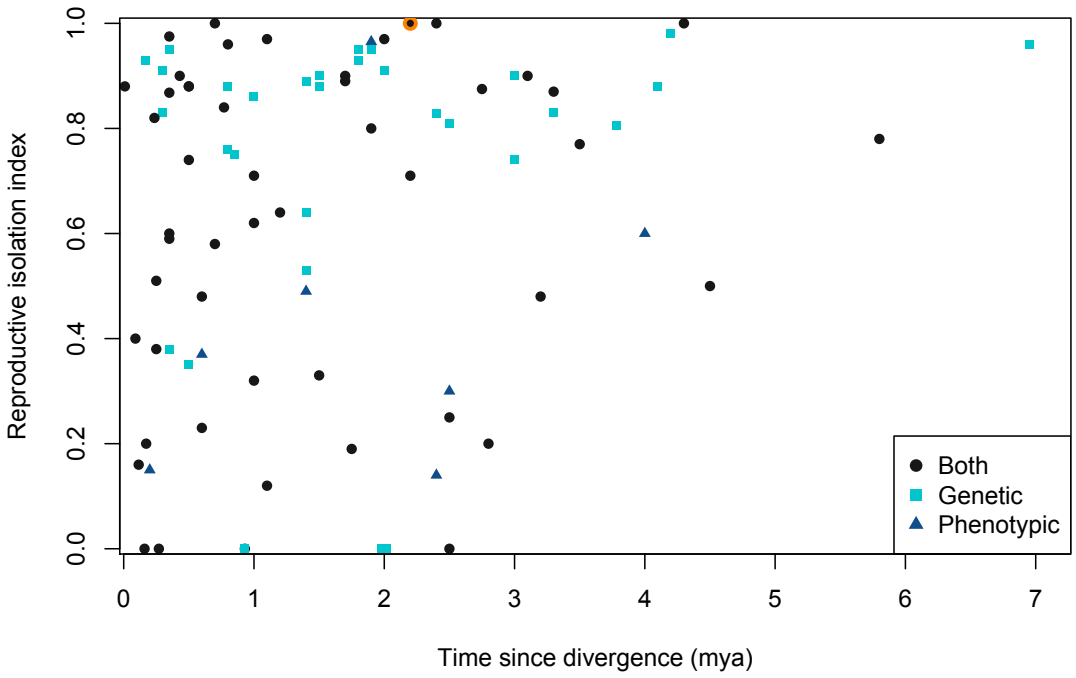

Figure 3.5: Gene flow across avian congeneric species and subspecies in secondary contact. A plot of the reproductive isolation index (based on phenotypic features and/or mitochondrial/genomic markers) versus divergence time (in millions of years ago) for 88 pairs of congeneric avian species and subspecies in secondary contact. Our study of *Z. kulambangrae* and *Z. murphyi* is highlighted in orange, and is the youngest reported example to date of complete reproductive isolation in birds (using genomic data) in which the two species' breeding ranges overlap (see text for details).

Table 3.1: Genomic and phenotypic patterns of secondary contact in two closely-related lineages

Scenario	Genetic Patterns	Phenotypic Patterns	Examples
1. No gene flow	Distinct mtDNA haplotypes and genomic lineages with no evidence of hybridization	Distinct species-specific phenotypes	Very limited: usually not examined in depth, particularly within a genomic context Example: This study
2. Reinforcement that leads to limited gene flow	Possible low levels of past genomic hybridization detected, or similar to Scenario 1: depends on initial levels of hybridization upon secondary contact	Distinct species-specific phenotypes; character displacement in sexually selected traits or response may be detectable within the contact zone	Difficult to show empirically, similar to Scenario 1 and likely understudied due to bias against testing species that appear reproductively isolated. Theoretical work has shown reinforcement is possible over a wide range of conditions (e.g. Liou and Price 1994, Servedio and Noor 2003) Example: Collared and Pied Flycatchers (Sætre and Sætre 2010)
3. Hybridization but with introgression of advantageous traits beyond the contact zone	A specific trait shows introgression in one species beyond the genomic contact/hybrid zone	A specific trait (genetic or phenotypic) shows introgression in one species beyond the genomic contact/hybrid zone, otherwise distinct species-specific phenotypes outside of the contact zone	Yes, but also must show trait confers a selective advantage in the introgressed population beyond the contact zone Example: Red-backed Fairy Wren subspecies (Baldassarre et al. 2014)
4. Stable Hybrid Zone	Hybrids with genomic admixture and possible mismatched mtDNA haplotypes found in center of the hybrid zone, with a decrease in admixed individuals with distance from the zone	Phenotypic hybrids found in the center of the hybrid zone, with a decrease in mixed phenotypic traits in individuals with distance from the zone	Yes, often stable hybrid zones are older in age. Selection against hybrids is counter-balanced with new dispersal into the zone Example: Yellow-rumped Warbler subspecies (Brelsford and Irwin 2009)
5. Fusion	Lineages have a similar genomic makeup but distinct mtDNA haplotypes still apparent	Fairly uniform phenotype across species range, is phenotypically classified as a single species throughout its range	Yes, several cases in which multiple mtDNA lineages have been found within a single species Example: Common Raven (Webb et al. 2011)
6. Hybrid Speciation	Hybrid lineage is a genomic mixture of two parental lineages and mtDNA haplotypes	Hybrid lineage is distinct and phenotypically diverged from the two parental lineages	Yes, several cases discovered recently with development of GBS and whole-genome sequencing Example: Golden-crowned Manakin (Barrera-Guzmzán et al. 2017)

Table 3.2: Mean ± SE for each of the six morphometric measurements taken from each captured bird of both species of *Zosterops* (K = *Z. kulambangrae* and M = *Z. murphyi*). P-values for t-tests have been Bonferroni corrected for multiple comparisons.

MEASUREMENT	Mean ± SE	t	p
Beak Length (mm)	K: 13.92 ± 0.09	-5.09	<0.001
	M: 14.54 ± 0.08		
Beak Width (mm)	K: 3.78 ± 0.04	-2.51	0.092
	M: 3.9 ± 0.03		
Beak Depth (mm)	K: 3.48 ± 0.02	-6.85	<0.001
	M: 3.73 ± 0.03		
Wing Chord (mm)	K: 61.01 ± 0.23	-8.46	<0.001
	M: 63.91 ± 0.25		
Tarsus (mm)	K: 17.57 ± 0.1	-5.69	<0.001
	M: 18.52 ± 0.14		
Tail Length (mm)	K: 38.63 ± 0.27	-7.4	<0.001
	M: 41.64 ± 0.3		
Mass (g)	K: 14.63 ± 0.1	-7.48	<0.001
	M: 15.9 ± 0.14		

Table 3.3: Studies in the review of reproductive isolation index versus divergence time (mya) for two congeneric species/subspecies in secondary contact with a known hybrid zone or parapatric distribution

System	Scientific Genus and Species	Divergence (mya)	RI Index	Method for RI Index	Category	Citation	Original Citation
Solomons White-eye and Kolombangara White-eye	*Zosterops kulambangrae* and *Z. murphyi*	2.2	1	Genomic SNPs, mtDNA, phenotyp	BOTH	This study	NA
Reunion Grey White-eye complex	*Zosterops borbonicus*	0.43	0.9	Plumage, AFLPs	BOTH	Warren et al. 2006 Molecular Ecology, Milá et al. 2010 BMC Evolutionary Biology	NA
Sooty Myzomela and Cardinal Myzomela	*Myzomela tristrami* and *M. cardinalis*	5.8	0.78	Genomic SNPs, mtDNA, plumage	BOTH	Sardell & Uy 2016 Evolution	NA
Rock Partridge and Red-legged Partridge	*Alectoris graeca* and *A. rufa*	2.4	0.14	Plumage	PHENOTYPIC	Price 2008 Appendix 15.1	Randi and Bernard-Laurent 1999 Auk, Randi 1996 Molecular Phylogenetics and Evolution
California Quail and Gambel's Quail	*Callipepla californica* and *C. gambelli*	1.1	0.12	Gen markers, morph, plumage	BOTH	Price 2008 Appendix 15.1	Zink and Blackwell 1998 Auk, Gee 2003 Evolution
Northern Flicker subspecies	*Colaptes auratus auratus* and *C. a. cafer*	0.115	0.16	Plumage, morph, GBS	BOTH	Price 2008 Appendix 15.1; Aguillon et. al 2018 Auk	Moore et al. 1991 Molecular Biology and Evolution, Moore 1987 Evolution
Western Gull and Glaucous-winged Gull	*Larus occidentalis* and *L. glaucescens*	0.6	0.23	Plumage, microsats	BOTH	Price 2008 Appendix 15.1	Gay et al. 2005 Auk, Bell 1996 Condor
Carrion Crow and Hooded Crow	*Corvus corone* and *C. cornix*	0.009	0.88	Plumage, mtDNA	BOTH	Price 2008 Appendix 15.1; Randler 2007 Ardea, Parkin et al. 2003 British Birds	Kryukov and Suzuki 2000 Russian Journal of Genetics, Saino & Villa 1992 Auk
Australian Magpie subspecies	*Gymnorhina tibicen*	0.09	0.4	Phenotype, mtDNA	BOTH	Price 2008 Appendix 15.1	Hughes et al. 2001 Biological Journal of the Linnean Society, Burton & Martin 1976 Emu
Collared Flycatcher and Pied Flycatcher	*Ficedula albicollis* and *F. hypoleuca*	2	0.97	Plumage, autosomal and Z-linked	BOTH	Price 2008 Appendix 15.1	Saetre et al. 2003 Proceedings of the Royal Society of London B, Veen et al. 2001 Nature
Black-capped Chickadee and Carolina Chickadee	*Poecile atricapillus* and *P. carolinensis*	1.7	0.89	mtDNA, nuclear, and Z-linked loci	BOTH	Price 2008 Appendix 15.1; Taylor et al. 2014 Current Biology	Sattler and Braun 2000 Auk, Bronson et al. 2005 Auk
Tufted Titmouse and Black-crested Titmouse	*Baeolophus bicolor* and *B. atricristatus*	0.2	0.15	Phenotype	PHENOTYPIC	Price 2008 Appendix 15.1	Klicka & Zink 1997 Science, Dixon 1955 University of California Publications in Zoology
African Red-eyed Bulbul and Garden/Common Bulbul	*Pycnonotus nigricans* and *P. barbatus*	2.5	0.3	Color of eye wattle	PHENOTYPIC	Price 2008 Appendix 15.1	Lloyd et al. 1997 Ostrich
Melodius Warbler and Icterine Warbler	*Hippolais polyglotta* and *H. icterina*	1.9	0.965	Song and morph	PHENOTYPIC	Price 2008 Appendix 15.1	Helbeig and Seibold 1999 Molecular Phylogenetics and Evolution, Faivre et al. 1999 Journal of Avian Biology, Secondi et al. 2003 Biologic
Common Chiffchaf and Iberian Chiffchaf	*Phylloscopus collybita* and *P. ibericus*	3.1	0.9	Song and AFLPs	BOTH	Price 2008 Appendix 15.1	Helbig et al. 2001 Journal of Evolutionary Biology, Bensch et al. 2002 Molecular Ecology
House Sparrow and Spanish/Willow Sparrow	*Passer domesticus* and *P. hispaniolensis*	4	0.6	Plumage	PHENOTYPIC	Price 2008 Appendix 15.1	Summers-Smith 1988 The Sparrows: A study of the genus Passer, Lockley 1992 Journal fur Ornithologie
Collared Towhee and Spotted Towhee	*Pipilo acai* and *P. maculatus*	2.5	0	Phenotype; GBS	BOTH	Price 2008 Appendix 15.1; Kingston et al. 2014 Ecology and Evolution, Kingston et al. 2017 Journal of Evolutionary Biology	Sibley 1950 University of California Publications in Zoology
Blue-winged Warbler and Golden-winged Warbler	*Vermivora cyanoptera* and *V. chrysoptera*	1.7	0.9	Phenotype, mtDNA	BOTH	Price 2008 Appendix 15.1; Shapiro et al. 2004 Auk	I. Lovette Personal Communication, Confer and Larkin 1998 Auk, Gill et al. 2001 The Birds of North America
Yellow-rumped Warbler subspecies	*Setophaga coronata coronata* and *S. c. auduboni*	0.16	0	Morph, allozymes, plumage, calls	BOTH	Price 2008 Appendix 15.1; Milá et al. 2007 Molecular Ecology, Brelsford and Irwin 2009 Evolution	Hubbard 1969 Auk
Hermit Warbler and Townsend's Warbler	*Setophaga occidentalis* and *S. townsendi*	0.5	0.74	Plumage, mtDNA	BOTH	Price 2008 Appendix 15.1	Rohwer et al. 2001 Evolution, Pearson 2000 Behavioral Ecology
Rose-breasted Grosbeak and Black-headed Grosbeak	*Pheucticus ludovicianus* and *P. melanocephalus*	2.2	0.71	Plumage, morph, mtDNA, nuclear	BOTH	Price 2008 Appendix 15.1; Mettler et al. 2009 Molecular Ecology	Klicka & Zink 1997 Science, West 1962 Auk
Indigo Bunting and Lazuli Bunting	*Passerina cyanea* and *P. amoena*	3.3	0.87	Song, morph, plumage, mtDNA, n	BOTH	Price 2008 Appendix 15.1; Carling and Brumfield 2008 Evolution	Baker & Johnson 1998 Auk, Emlen et al. 1975 Wilson Bulletin, Baker & Boylan 1999 Condor, Kroodsma 1975 Auk
Vinaceous Dove and Ring-necked Dove	*Streptopelia vinacea* and *S. capicola*	1.5	0.33	Plumage, song, mtDNA, AFLPs	BOTH	Price 2008 Appendix 15.1; den Hartog et al. 2010 Auk	NA
Variable Antshrike complex	*Thamnophilus caerulescens*	0.93	0	Plumage, song, mtDNA	BOTH	Price 2008 Appendix 15.1	Brumfield 2005 Auk
Yellowhammer and Pine Bunting	*Emberiza citrinella* and *E. leucocephalos*	4.5	0.5	AFLPs, mtDNA, phenotype	BOTH	Price 2008 Appendix 15.1; Irwin et al. 2009 Biological Journal of the Linnean Society	Panov et al. 2003 Dutch Birding
Baltimore Oriole and Bullock's Oriole	*Icterus galbula* and *I. bullockii*	3.2	0.48	Plumage, mtDNA, nuclear marker	BOTH	Price 2008 Appendix 15.1; Carling et al. 2011 Auk	Rising 1996 Condon, Allen 2002 Ph.D. Thesis Univeristy of Indiana-Bloomington, Edinger 1985 M.S. Thesis-University of Minnesota
Red-backed Fairy Wren subspecies	*Malurus melanocephalus cruentatus* and *M. m. melanocephalu*	0.27	0	Genomic SNPs, Plumage	BOTH	Lee & Edwards 2008 Evolution, Baldassarre et al. 2014 Evolution	NA
Blue-footed Booby and Peruvian Booby	*Sula nebouxii* and *S. variegata*	0.35	0.975	Microsats, mtDNA, phenotype	BOTH	Friesen & Anderson 1997 Molecular Phylogenetics and Evolution, Patterson et al. 2010 Molecular Phylogenetics and Evolution, Taylor et al. 2010 Waterbirds, Taylor et al. 2012 Journa	NA
MacGillivray's Warbler and Mourning Warbler	*Geothlypis tolmiei* and *G. philadelphia*	1.2	0.64	Phenotype, mtDNA, Z-linked mar	BOTH	Weir & Schluter 2004 Proceedings of the Royal Society London B, Irwin et al. 2009 Journal of Avian Biology	NA
Greenish Warbler ring species complex	*Phylloscopus trichloloides viridanus* and *P. t. plumbeitarsus*	1.5	0.9	Genomic SNPs	GENETIC	Price 2008, Alcaide et al. 2014 Nature	NA
Cape White-eye subspecies	*Zosterops capensis* and *Z. v. pallidus*	0.77	0.84	Phenotype, mtDNA, microsats	BOTH	Oatley et al. 2012 Molecular Phylogenetics and Evolution, Oatley et al. 2017 Biological Journal of the Linnean Society	NA
Townsend's Warbler and Black-throated Green Warbler	*Setophaga townsendi* and *S. virens*	1	0.62	Plumage, morph, mtDNA, two ger	BOTH	Toews et al. 2011 Journal of Avian Biology, Lovette et al. 2010 Molecular Phylogenetics and Evolution	NA
Pacific Wren and Winter wren	*Troglodytes pacificus* and *T. hiemalis*	4.3	1	AFLPs, mtDNA, song, morph	BOTH	Toews and Irwin 2008 Molecular Ecology	NA
Red-breasted Sapsucker and Yellow-bellied Sapsucker	*Sphyrapicus ruber* and *S. varius*	1.1	0.97	Plumage, morph, GBS	BOTH	Seneviratne et al. 2016 Journal of Avian Biology, Cicero and Johnson 1995 Auk, Weir and Schulter 2004 Proceedings of the Royal Society London B	NA
Red-breasted sapsucker and Red-naped Sapsucker	*Sphyrapicus ruber* and *S. nuchalis*	0.5	0.88	Plumage, morph, GBS	BOTH	Seneviratne et al. 2016 Journal of Avian Biology, Cicero and Johnson 1995 Auk, Weir and Schulter 2004 Proceedings of the Royal Society London B; Also in Price 2008 Appendix 15.1	NA
Pacific-slope Flycatcher and the Cordilleran Flycatcher	*Empidonax difficilis* and *E. occidentalis*	0.35	0.38	mtDNA, AFLPs	GENETIC	Rush et al. 2009 Journal of Avian Biology, Johnon and Cicero 2002 Molecular Ecology	NA
Willet subspecies	*Tringa semipalmata semipalmata* and *T. s. inornata*	0.7	1	Morph, mtDNA, UCEs	BOTH	Oswald et al. 2016 Auk	NA
Black-throated Finch subspecies	*Poephila cincta atropygialis* and *P. c. cincta*	0.6	0.37	Morph	PHENOTYPIC	Price 2008 Appendix 15.1; Jennings and Edwards 2005 Evolution	Ford 1985 Emu
White-collared Manakin and Golden-collared Manakin	*Manacus candei* and *M. vitellinus*	2.8	0.2	Plumage, genetic markers, GBS	BOTH	Price 2008 Appendix 15.1; Parchman et al. 2013 Molecular Ecology, Brumfield et al. 2008 Systematic Biology	Brumfield et al. 2001 Evolution
Northern Jacana and Wattled Jacana	*Jacana spinosa* and *J. jacana*	0.7	0.58	Phenotype and mtDNA	BOTH	Miller et al. 2014 BMC Evolutionary Biology	NA
Common Redstart (distinct mtDNA lineages)	*Phoenicurus phoenicurus*	2	0	mtDNA	GENETIC	Hogner et al. 2012 Ecology and Evolution	NA
Common Raven (distinct mtDNA lineages)	*Corvus corax*	2	0	mtDNA	GENETIC	Webb et al. 2011 Molecular Ecology	NA
Hume's Leaf Warbler and Yellow-browed Warbler	*Phylloscopus humei* and *P. inornatus*	2.4	1	Song, mtDNA	BOTH	Irwin et al. 2001 Ibis	NA
Western Scrub Jay subspecies	*Aphelocoma californica superciliosa* and *A. c. nevadae*	2.5	0.25	Microsats, mtDNA, phenotype	BOTH	Gowen et al. 2014 BMC Evolutionary Biology	NA
Swainson's Thrush subspecies	*Catharus ustulatus ustulatus* and *C. u. swainsoni*	0.35	0.6	mtDNA, AFLPs, morph, plumage	BOTH	Ruegg 2007 Ornithological Monographs, Ruegg 2008 Evolution	NA
Barn Swallow subspecies	*Hirundo rustica tytleri* and *H. r. gutturalis*	0.25	0.51	plumage, morph, GBS	BOTH	Scordato et al. 2017 Molecular Ecology	NA
Barn Swallow subspecies	*Hirundo rustica rustica* and *H. r. tytleri*	0.25	0.38	plumage, morph, GBS	BOTH	Scordato et al. 2017 Molecular Ecology	NA
Lucifer Hummingbird migratory and sedentary populations	*Calothorax lucifer*	0.3	0.83	mtDNA, microsats	GENETIC	Licona-Vera et al. 2018 Biological Journal of the Linnean Society	NA
Saltmarsh Sparrow and Nelson's Sparrow	*Ammodramus caudacutus* and *A. nelsoni*	0.6	0.48	mtDNA, microsats, morph, pluma	BOTH	Walsh et al. 2015 Auk	NA
Woodpecker Finch and Mangrove Finch	*Camarhynchus pallidus* and *C. heliobates*	0.171	0.93	Microsats	GENETIC	Lawson et al. 2017 Conservation Genetics, Lamichhaney et al. 2015 Nature	NA
Lesser Scaup and Greater Scaup	*Aythya affinis* and *A. marila*	0.35	0.95	mtDNA, ddRAD	GENETIC	Lavretsky et al. 2016 Molecular Ecology	NA
Western Wood-Pewee and Eastern Wood-Pewee	*Contopus sordidulus* and *C. virens*	0.85	0.75	GBS	GENETIC	Manthey and Robbins 2016 Avian Research, Cicero and Johnson 2002 Mol Phylogenetics and Evolution	NA
Common Murre and Thick-billed Murre	*Uria aalge* and *U. lomvia*	4.2	0.98	mtDNA and nuclear introns	GENETIC	Taylor et al. 2012 Polar Biology	NA
Great Reed Warbler and Clamorous Reed Warbler	*Acrocephalus arundinaceus* and *A. stentoreus*	2	0.91	Microsats	GENETIC	Hansson et al. 2012 PLoS ONE	NA
Lowland Tiny Greenbul subspecies (lowland and montane)	*Phyllastrephus debilis rabai* and *P. d. albigula*	2.75	0.875	Morph, plumage, mtDNA, introns	BOTH	Fuchs et al. 2011 BMC Evolutionary Biology	NA
Savi's Warbler (distinct mtDNA lineages)	*Locustella luscinioides*	0.93	0	mtDNA, microsats	GENETIC	Neto et al. 2012 PLoS ONE	NA
Grey-breasted Wood-wren subspecies (low and high elevation)	*Henicorhina leucophrys hilaris* and *H. l. leucophrys*	3.78	0.806	Song, AFLPs	GENETIC	Halfwerk et al. 2016 Journal of Evolutionary Biology, Dingle et al. 2006 Auk, Dingle et al. 2008 Journal of Evolutionary Biology	NA
Plain-brown Woodcreeper subspecies	*Dendrocincla fuliginosa atrirostris* and *D. f. rufoolivacea*	3	0.74	mtDNA and GBS	GENETIC	Weir et al. 2015 Evolution	NA
Elegant Woodcreeper and Spix's Woodcreeper	*Xiphorhynchus elegans elegans* and *X. spixii*	2.5	0.81	mtDNA and GBS	GENETIC	Weir et al. 2015 Evolution	NA
Wedge-billed Woodcreeper subspecies	*Glyphorynchus spirurus inornatus* and *G. s. paraensis*	3.3	0.83	mtDNA and GBS	GENETIC	Weir et al. 2015 Evolution	NA
Rondonia Warbling Antbird and Spix's Warbling Antbird	*Hypocnemis ochrogyna* and *H. striata striata*	1.9	0.95	mtDNA and GBS	GENETIC	Weir et al. 2015 Evolution	NA
White-breasted Antbird and Bare-eyed Antbird	*Rhegmatorhina hoffmannsi* and *R. gymnops*	0.8	0.88	mtDNA and GBS	GENETIC	Weir et al. 2015 Evolution	NA
Snow-capped Manakin and Opal-crowned Manakin	*Lepidothrix nattereri* and *L. iris eucephala*	0.8	0.76	mtDNA and GBS	GENETIC	Weir et al. 2015 Evolution	NA
Common Scale-backed Antbird and Xingu Scale-backed Antbird	*Willisornis poecilinotus griseiventris* and *W. vidua nigrigula*	4.1	0.88	mtDNA and GBS	GENETIC	Weir et al. 2015 Evolution	NA
Black Stilt and Pied Stilt	*Himantopus novaezelandiae* and *H. himantopus leucocephalus*	1	0.71	mtDNA and microsats, plumage	BOTH	Steeves et al. 2010 Molecular Ecology; Wallis 1999 Conservation Advisory Science Notes No. 239	NA
Green-fronted Hummingbird and Violet-crowned hummingbird	*Amazilia viridifrons* and *A. violiceps*	0.1725	0.2	Morph and microsats	BOTH	Rodriguez and Ornelas 2018 Journal of Avian Biology	NA
Balearic Shearwater and Yelkouan Shearwater	*Puffinis mauretanicus* and *P. yelkouan*	1	0.32	Phenotype and microsats	BOTH	Genovart et al. 2007 Biological Conservation, Genovart et al. 2012 Journal of Heredity	NA
Trindade Petrel and Kermadec Petrel	*Pterodroma arminjoniana* and *P. neglecta*	0.5	0.88	Phenotype, microsats, and song	BOTH	Brown et al. 2010 Molecular Ecology; Brown et al. 2011 PLoS ONE	NA
Canada Goose and Cackling Goose	*Branta canadensis* and *B. hutchinsii*	3.5	0.77	Morph, mtDNA, microsats	BOTH	Leafloor et al. 2013 AUK	NA
Willow Grouse and Rock Ptarmigan	*Lagopus lagopus* and *L. muta*	0.8	0.96	mtDNA, microsats, plumage	BOTH	Quintela et al. 2010 Conservation Genetics	NA
Chukar Partridge and Red-legged Partridge	*Alectoris chukar* and *A. rufa*	2.4	0.828	mtDNA, microsats	GENETIC	Tejedor et al. 2007 Journal of Heredity, Randi 1996 Molecular Phylogenetics and Evolution	NA
Greater Spotted Eagle and Lesser Spotted Eagle	*Aquila clanga* and *A. pomarina*	1.5	0.88	GBS and microsats	GENETIC	Vali et al. 2010 Biological Journal of the Linnean Society, Helbig et al. 2005 Journal of Ornithology	NA
Thrush Nightingale and Common Nightingale	*Luscinia luscinia* and *L. megarhynchos*	1.8	0.95	Autosomal and Z-linked introns (s	GENETIC	Storchova et al. 2010 Evolution	NA
Hawaiian Duck and North American Mallard	*Anas wyvilliana* and *A. platyrhynchos*	1.4	0.53	mtDNA, microsats, AFLPs	GENETIC	Fowler et al. 2008 Conservation Genetics	NA
Herring Gull and Glaucous Gull	*Larus argentatus* and *L. hyperboreus*	0.35	0.868	Morph, mtDNA, microsats	BOTH	Palsson et al. 2009 Auk	NA
Herring Gull and Caspian Gull	*Larus argentatus* and *L. cachinnans*	0.35	0.59	Phenotype, mtDNA, microsats	BOTH	Gay et al. 2007 Molecular Ecology	NA
Chatham Island Red-crowned Parakeet and Forbes' Parakeet	*Cyanoramphus novaezelandiae* and *C. forbesi*	1.75	0.19	Morph, mtDNA, microsats	BOTH	Boon et al. 2000 Bird Conservation International, Chan et al. 2006 Conservation Genetics	NA
Thick-billed Fox Sparrow, Slate-colored Fox Sparrow, and Red Fox Sp	*Passerella megarhyncha* , *P. schistacea* , and *P. iliaca*	0.5	0.35	mtDNA	GENETIC	Zink 1994 Evolution	NA
Mallard and Grey Duck	*Anas platyrhynchos* and *A. superciliosa*	1.4	0.49	Morph	PHENOTYPIC	Gillespie 1985 Auk, Lavretsky et al. 2014 Molecular Phylogenetics and Evolution	NA
American Golden-Plover and Pacific Golden-Plover	*Pluvialis dominica* and *P. fulva*	1.8	0.93	mtDNA and AFLPs	GENETIC	Withrow and Winkler 2014 Wilson Journal of Ornithology	NA
Cactus Finch and Medium Ground Finch	*Geospiza fortis* and *G. scandens*	0.236	0.82	Morph, mtDNA, microsats	BOTH	Grant et al. 2004 Evolution, Lamichhaney et al. 2014 Nature	NA
Oak Titmouse and Juniper Titmouse	*Baeolophus inornatus* and *B. ridgwayi*	1.9	0.8	Morph, mtDNA, allozymes	BOTH	Cicero 2004 Evolution	NA
Common Wild Quail and domesticated Japanese Quail	*Coturnix coturnix coturnix* and *C. japonica*	0.3	0.91	mtDNA and microsats	GENETIC	Barilani et al. 2005 Biological Conservation, Seabrook-Davidson et al. 2009 PLoS One, see also Chazara et al. 2010 Conservation Genetics	NA
Mallard and Mottled Duck	*Anas platyrhynchos* and *A. fulvigula*	1.4	0.89	Microsats	GENETIC	Williams et al. 2005 Conservation Genetics, Lavretsky et al. 2014 Molecular Phylogenetics and Evolution	NA
Mallard and Black Duck	*Anas platyrhynchos* and *A. rubripes*	1.4	0.64	Microsats	GENETIC	Mank et al. 2004 Conservation Genetics	NA
Barred Owl and Northern Spottled Owl	*Strix varia* and *S. occidentalis caurina*	6.95	0.96	mtDNA and AFLPs	GENETIC	Haig et al. 2004 Conservation Biology	NA
Chinese Crested Tern and Greater Crested Tern	*Thalasseus bernsteini* and *T. bergii*	1	0.86	mtDNA and two Z-linked genes	GENETIC	Yang et al. 2018 IBIS	NA
Great Tit and Japanese Tit	*Parus major* and *P. minor*	3	0.9	mtDNA	GENETIC	Federov et al. 2006 Zoologicheskii Zhurnal, Kvist et al. 2003 Proceedings of the Royal Society London B	NA

Table 3.4: Estimates of time as a predictor of reproductive isolation in logistic regression models

Logistic Model	**Est (SE)**	**Z**	**p**
All data (N = 88)	0.22(0.18)	1.24	0.22
Both phen and gen (N = 50)	0.19(0.24)	0.81	0.42
Pheno est only (N = 7)	0.28(0.66)	0.42	0.68
Gen est only (N = 31)	0.25(0.33)	0.76	0.45
Forced Intercept Logistic Model	**Est (SE)**	**Z**	**p**
All data (N = 88)	0.32(0.12)	2.69	0.007
Both phen and gen (N = 50)	0.27 (0.16)	1.66	0.098
Pheno est only (N = 7)	-0.03(0.34)	-0.083	0.934
Gen est only (N = 31)	0.50(0.22)	2.27	0.023

Chapter 4

Small estimated population sizes of the two Kolombangara White-eye *Zosterops* species suggest need for stringent conservation efforts

Summary

It is essential to study endemic populations of animals on small islands to understand the current population size, range, and genetic structuring, which in turn allows for the maximization of conservation efforts and a mitigation of human impacts. Two *Zosterops* White-eye bird species live on the small island of Kolombangara in the Solomon Islands, one of which is endemic to high elevations on the island (*Z. murphyi*), while the other is also found on the large neighboring island of New Georgia and its smaller surrounding islands (*Z. kulambangrae*). We estimated the current-day effective population sizes for both species using two methods designed to work with genome-by-sequencing single nucleotide polymorphism (SNP) data. We also examined genetic structuring within species. We found that both species have similar small effective population sizes estimated at between approximately 400-2500 individuals each (using different computational methods), and that within the small island, little to no genetic structuring is apparent in either species. Since the geographic range of *Z. kulambangrae* encompasses more than half of the islands in the New Georgia Province, the low estimated effective population size for this species is especially concerning. Overall, given the extremely limited effective population sizes of both species, it is necessary to conserve the remaining forest habitat available on Kolombangara and the neighboring New Georgia islands and to prevent additional logging and destruction of the natural tropical forest habitat.

Background

The effects of the Anthropocene on biodiversity are unprecedented (Dirzo et al. 2014, Johnson et al. 2017). In addition to climate change drastically affecting habitats, other human-induced effects such as deforestation, pollution, introduction of exotic species, and alteration of habitats are rapidly affecting organisms around the world. Particularly at risk are organisms endemic to small ranges in very specific habitat types (e.g. Johnson and Stattersfield 1990, Benning et al. 2002, Williams et al. 2003, Loarie et al. 2008, Dirnböck et al. 2011, Urban 2015, Allen and Lendemer 2016). It is imperative that we fully understand the biology of these endemic organisms, as such an understanding would allow for the maximization of conservation efforts and a mitigation of human impacts.

The Solomon Islands in the South Pacific are an incredibly biodiverse chain of islands. Densely covered in wet tropical rain forest (Whitmore 1969), there is an extremely high diversity of organisms found on this island archipelago, particularly in frogs, bats, rodents, and birds (Mayr and Diamond 2001). The island of Kolombangara is a small, extinct stratovolcano island located in the New Georgia (Western) Province of the Solomon Islands. The name Kolombangara translates to “Water King,” which is apt, as the island receives over 3000 mm of precipitation per year. At 688-square kilometers in size and 1770 meters in elevation, it is one of the most biodiverse islands in the region with forest habitat from the coast to the crater rim (Whitmore 1969). For example, 82 out of 85 total bird species found in the New Georgia Province can be found on the island of Kolombangara (Mayr and Diamond 2001). However, 90% of the land below 400 meters

has been previously logged, turned into plantation forest, or cleared for human use (Katovi et al. 2012, 2015, 2016).

Two species of *Zosterops* White-eyes are found on Kolombangara island: *Z. murphyi* and *Z. kulambangrae* (Fig. 4.1). These two species are part of the Zosteropidae family of birds, which is one of the most speciose vertebrate families (and known as a "great speciator" (see Diamond et al. 1976)), with a large portion of species as island endemics (van Balen 2001, Moyle et al. 2009). *Z. murphyi* is endemic to the forest in the upper elevations (600m and above, Weeks et al. 2016, Cowles & Uy 2019) on the island of Kolombangara. *Z. kulambangrae* has a broader geographic distribution with populations on the nearby large island of New Georgia, the smaller islands of Vangunu, Gatokae, Vonavona, Kohinggo (also known as Arundel), and all of their nearby satellites (Dutson 2011, see Fig. 4.1), which, in addition to the islands of Tetepare and Rendova (each with their own species of White-eye), were likely joined into a single land mass during the late Pleistocene (named "Greater Gatumbangra") (Mayr and Diamond 2001, Weeks et al. 2016). Particularly on Kolombangara, *Z. kulambangrae* can be found in patches of tropical forest as low as 50 meters on the island, and up to around 1000 meters in elevation on the forest slopes (Cowles and Uy 2019).

Surprisingly, no gene flow or hybridization has been detected between *Z. kulambangrae* and *Z. murpyhi* on Kolombangara, even though they associate in flocks together between 600-1000 meters in elevation on the island, and have only been diverging for approximately two million years (Cowles and Uy 2019). In terms of conservation, this lack of gene flow is positive, as the endemic *Z. murphyi* remains a distinct species and does not suffer the effects of external gene flow. Gene flow into

endemic species on small islands can sometimes contribute to extinction of the endemic taxa (e.g. Genovart et al. 2007, Lawson et al. 2017), so it is advantageous these two species show complete reproductive isolation and can be considered two distinct species units.

From a conservation perspective, almost nothing is known about the population sizes of these two species or how historical and present-day logging on Kolombangara has affected their numbers. Mayr and Diamond (2001) roughly estimated that these two bird species fell into their category of "abundant," which they defined as >100 pairs per km^2. However, we can use genomic data to better estimate effective population size. Compared to the actual population census size (N), the effective population size (N_e) of a population is defined as the size of an idealized Wright-Fisher population (a population of diploid individuals with no migration, selection, and equal mating success among members), which would give rise to the same rate of inbreeding and change in gene frequencies per generation (i.e., loss of genetic diversity) as observed in the actual population (Fisher 1930, Wright 1931, Crow and Kimura 1970, Wang et al. 2016). Several methods have been developed to calculate N_e using genomic data. These methods include heterozygote excess, coalescent-based inference, linkage disequilibrium, and temporal sampling (see Wang 2005, Wang et al. 2016 for a review). Some methods such as coalescence using the program PSMC (Li and Durbin 2011) require high genome-wide coverage and little missing data to make accurate estimates of N_e (e.g. see Nadachowska-Brzyska et al. 2016 for an example using the program PSMC in *Ficedula* flycatchers). However, a few programs have been developed to work with reduced representation

genome-by-sequencing (GBS) data, which consist of thousands of single nucleotide polymorphisms (SNPs) scattered throughout the genome.

The aim of our study was to estimate the present-day effective population sizes of the two *Zosterops* White-eye species found on the island of Kolombangara, and to examine fine-scale population structuring. We used two computational programs (one coalescence-based and the other linkage disequilibrium-based) that can use single time point samples and GBS SNP data to estimate present-day effective population sizes. We also examined genetic relatedness and population structuring between individuals using two programs (fineRADstructrue and SPAGeDi). By providing current population size estimates, we can offer recommendations to maximize the conservation efforts for these two species.

Methods

Genomic Sequencing

We collected blood samples from a total of 76 *Z. kulambangrae* and 58 *Z. murphyi* that were captured along two distinct elevational transects (Iriri Corridor and Ringgi Mt. Veve Trail from Imbu Rano) located on different sides of Kolombangara Island during our 2016 and 2017 field seasons (see Cowles and Uy 2019 for further sampling details). Of the 76 *Z. kulambangrae*, 60 were male given that we mostly caught territorial individuals (sometimes pairs) from this species using target-netting and song playback. Of the 58 *Z. murphyi*, we obtained a more even sex ratio due to passive netting of non-territorial birds, as 30 of the birds were male. Although we obtained different sex ratios in our sampling across species, we do not believe this capturing bias is reflective of

the actual population sex ratio or that it has an effect on the estimates of effective population size given we used genome-wide datasets. DNA was extracted using Qiagen's DNeasy Blood and Tissue Extraction kits (Qiagen, Hilden, Germany), and sent to the University of Wisconsin-Madison's Bioinformatics Resource Center for genotyping-by-sequencing using the restriction enzyme ApeKI.

SNPs were called using the TASSEL pipeline and parameters (Glaubitz et al. 2014) and mapped to the Zebrafinch (*Taeniopygia guttata*) ENSEMBL genome assembly (taeGut3.2.4.87). After filtering (see Cowles and Uy 2019 for further details), we ended up with a dataset of 79,742 mapped SNPs. For our subsequent population size analyses, we used a reduced dataset of 23,752 SNPs that had a minor allele frequency of 0.05 and were found in at least 70 of the 134 sequenced individuals. Finally, we created a vcf file with these 23,752 SNPs in TASSEL version 1.5 (Bradbury et al. 2007) for use with our population size analyses.

Population Size Estimates

Stairway Plot

To estimate contemporary effective population sizes (N_e) for both species, we first used a program called 'Stairway Plot v0.2' (Liu and Fu 2015). Stairway Plot is a program that estimates population sizes over time using SNP frequency spectra and a flexible multi-epoch model (Liu and Fu 2015). In the program, each SNP is treated as an independent unlinked locus, and coalescent trees from each SNP are used to calculate a composite likelihood of the given site frequency spectrum (SFS), which is the distribution of derived allele frequencies in a given population. Finally, the estimation of

the composite likelihood and a user-input mutation rate are used to calculate the current number of individuals in the population over time (see Liu and Fu (2015) for full mathematical details).

To use the program, we first created a folded SFS file for each species from our vcf file using the 'vcf2sfs' R scripts provided in Liu et al. (2018). A folded SFS is a histogram of minor allele frequencies across polymorphic sites. Next, we used the program Stairway Plot v0.2 implemented in Java to estimate the contemporary N_e for each folded SFS file. In comparison to other methods used to estimate N_e, this program is computationally efficient, works well with larger sample sizes, and does not need a pre-defined demographic model (Liu and Fu 2015). Although Stairway Plot has not been fully tested for SNP datasets that include some missing data, we assume the SFS generated from our GBS SNP dataset is similar in shape to the real SFS (without missing data). In Stairway Plot, we first created input blueprint files. For each *Zosterops* species, we assumed a generation time of one year (*Zosterops* species likely have short generation times: see Moyle et al. 2009), and used an overall genomic mutation rate per generation of 4.6×10^{-9}, a rate consistent with estimates of mutation rates in other bird lineages (e.g. Helm-Bychowski and Wilson 1986, Axelsson et al. 2004, Nam et al. 2010, Smeds et al. 2016). Although we do not know the specific generation time or lineage-specific mutation rate for these two *Zosterops* species, our overall goal was to obtain current population size estimates for the two species present on the island, which our estimated parameters allowed us to do. In the blueprint file, we used the suggested percentage of sites used for training (0.67), and the number of suggested random break points (Kulambangrae: 37 ,75, 113, 150; Murphyi: 29, 57, 86, 114) for each analysis. We used

two as the smallest SFS bin size (ignoring singletons due to filtering; results have been shown to be robust using this method (see Huang 2019)), and created 100 bootstrapped SFS files (number of iterations) for each model. The summary output model gives an estimated population size for the present day based on a summary of all the model iterations.

N_eEstimator

Our second method of assessing contemporary effective population size was done using the software program N_eEstimator version 2.0 (Do et al. 2014). N_eEstimator is a program that is able to produce estimates of contemporary population size using genepop or fstat input file format. We first converted our vcf file containing 23,752 SNPs into the fstat file format using the program PGDspider version 2.1.1.5 (Lischer and Excoffier 2011). We were then able to put our file into the N_eEstimator program, and used the linkage disequilibrium method, which is based on Waples and Do (2008) and has been used in other White-eye species (e.g. Husemann et al. 2015). Overall, increases in population size would result in a decrease in the expected levels of linkage disequilibrium throughout the genome across the population. The linkage disequilibrium method of Waples and Do (2008) assumes each SNP is an independent locus. Using the Burrows' delta method, a correlation coefficient is calculated for each pair of alleles at every SNP. This correlation coefficient is then averaged across all SNPs and used to estimate N_e, given the average correlation coefficient's expected value is derived from N_e, the number of individuals sampled, the recombination rate, and the mating system (see Waples and Do (2008) for complete mathematical details). Although N_eEstimator has two additional

possible methods for calculating effective population sizes (heterozygote excess and molecular coancestry), testing of the program by the authors showed that the linkage disequilibrium method was by far the most accurate method for single time point sampling of a population. We used a random mating linkage disequilibrium model, with a critical value of 0.05 (i.e., minimum allele frequency).

We also tested whether using less stringent filtering for minor allele frequency on the SNP dataset (i.e., different critical values) might change the population size estimates (see Nunziata and Weisrock 2018). We tested the N_eEstimator program using critical values of 0.02, 0.01, and 0 on SNP datasets unfiltered for minor allele frequency for each species (only filtered so that each SNP was present in at least half of the individuals per species). These datasets ended up being 31,132 SNPs for *Z. kulambangrae* and 30,539 SNPs for *Z. murphyi*.

Calculation of Coancestry and Relatedness Within Species

We used the program fineRADstructure (Malinsky et al. 2018) to assess fine-scale population and clustering within each species. This program calculates coancestry values between individuals in a pairwise manner (using RADpainter), and then clusters individuals and assesses population structure using a Markov chain Monte Carlo (MCMC) algorithm (using fineSTRUCTURE). The input for the program is a haplotype file, and we converted our vcf file of 23,752 SNPs into this file format using the script 'hapsFromVCF' in RADpainter. We then used RADpainter to calculate the coancestry values across all individuals, and then used fineSTRUCTURE to cluster individuals, using the input arguments of x 100,000, -z 100,000, and -y 1,000. Finally, we used

fineSTRUCTURE's tree building algorithm with the input arguments of -m T -x 10000 to create a simple tree of the individuals. We visualized the program output with the fineSTRUCTURE GUI (https://people.maths.bris.ac.uk/~madjl/finestructure/finestructure.html).

Additionally, we used the program SPAGeDi version 1.5 (Hardy and Vekemans 2002) to calculate relatedness estimates for both *Zosterops* species. We used our data set of 23,752 SNPs. We first calculated pairwise kinship coefficient estimates (Loiselle et al. 1995) among pairs of individuals within each species, and then calculated the means of these values for the two species.

Tajima's D

As a potential indicator of population size change, we calculated the statistic Tajima's D (Tajima 1989a, 1989b) for both species. When the D-statistic is equal to 0, the population is at a drift-mutation equilibrium, which is indicative of a stable population size. If Tajima's D is less than zero, it implies an excess of low frequency polymorphisms in the population and is an indication of a selective sweep or population expansion after a recent bottleneck. Finally, if Tajima's D is greater than zero, it means there are fewer low frequency polymorphisms in the population than expected, consistent with either balancing selection or population contraction during an immediate bottleneck. We first turned our vcf file of 23,752 SNPs into a "DNAbin" object in R using the 'vcf2DNAbin' script in the R package 'vcfR' (Knaus and Grunwald 2017). We then ran a Tajima's D test for both species separately, using the 'tajima.test' function in the R package 'pegas' (Paradis 2010). We also conducted Tajima's D tests on our vcf files

unfiltered for minor allele frequency for both species (31,132 SNPs for *Z. kulambangrae* and 30,539 SNPs for *Z. murphy*) as a check to confirm filtering did not affect our results.

Results

Estimates of contemporary population size

Using our bird-specific mutation rate of 4.6×10^{-9} per site per generation and an estimated generation time of one year, current effective population sizes using the Stairway Plot program are estimated at approximately 2,747 individuals (95% CI: 2616-6403 individuals) for *Z. kulambangrae* (found on Kolombangara and the neighboring large island of New Georgia, the smaller islands of Vangunu, Gatokae, Vonavona, Kohinggo/Arundel, and all of their nearby satellites) and approximately 2,636 individuals (95% CI: 1375-8195 individuals) for *Z. murphyi* (endemic to the island of Kolombangara) (Fig. 4.2).

Current effective population sizes for both species were also estimated using the program N_eEstimator (Fig. 4.2). For *Z. kulambangrae*, N_e was estimated at 673.4 (CI: 665.1-681.9) individuals and for *Z. murphyi*, N_e was estimated at 414.2 (CI: 409.2-419.4) individuals at a critical value (i.e., minimum allele frequency) of 0.05 on our original filtered dataset. To verify our results, we did additional model runs on the SNP datasets that were unfiltered for minor allele frequency before being input into the program. For *Z. kulambangrae,* we found that for a critical value of 0.05, N_e was estimated at 728.6 (CI: 719.7-737.7) individuals, for a critical value of 0.02, N_e was estimated at 810.7 (CI: 800.4-821.3) individuals, for a critical value of 0.01, N_e was estimated at 855.4 (CI: 844.1-867.1) individuals, and finally for a critical value of 0, N_e was estimated at 861.2

(CI: 849.8.7-872.9) individuals. For *Z. murphyi,* we found that for a critical value of 0.05, N_e was estimated at 478.9 (CI: 473.1-484.8) individuals, for a critical value of 0.02, N_e was estimated at 557.0 (CI: 549.7-564.5) individuals, for a critical value of 0.01, N_e was estimated at 628.5 (CI: 619.5-637.8) individuals, and finally for a critical value of 0, N_e was estimated at 631.5 (CI: 622.3-640.8) individuals. Overall, the estimated population sizes using the differently-filtered datasets were fairly similar in magnitude (i.e., the mid-hundreds) compared to the original datasets for both species.

Estimates of coancestry and relatedness among individuals

Overall higher relatedness among *Z. murphyi* is apparent in the fineRADstructure plot (Fig. 4.3), as pairs of *Z. murphyi* individuals share overall higher estimated co-ancestry levels compared to pairs of *Z. kulambangrae* individuals. The fineRADstructure algorithms did not cluster individuals within each species either by transect or by elevation (Fig. 4.3). Similar to Cowles and Uy (2019), fineRADstructure also clustered each species as a distinct group of individuals with no suggestion of gene flow (Fig. 4.3).

We also calculated pairwise kinship coefficient estimates among pairs of individuals within each species using the program SPAGeDi. *Z. murphyi* individuals had higher kinship coefficient estimates overall (mean $\pm$ SE = 0.145 $\pm$ 0.001) compared with *Z. kulambangrae* individuals (mean $\pm$ SE = 0.091 $\pm$ 0.001).

Tajima's D

For *Z. kulambangrae*, Tajima's D was negative and significantly different from 0 ($D = -3.17$, $p < 0.002$). For *Z. murphyi*, Tajima's D was also negative and significantly

different from 0 (D = -3.36, $p < 0.001$). Conducting Tajima's D tests on the datasets unfiltered for minor allele frequency gave similar results (*Z. kulambangrae*: D = -3.3, $p < 0.001$, *Z. murphyi* : D = -3.37, $p < 0.001$).

Discussion

Low N_e Sizes suggest need for strict conservation measures

Although we obtained different estimates for the effective population sizes of *Z. kulambangrae* and *Z. murphyi* using different genomic methods and filtering on datasets, all of our estimates point to effective population sizes in the mid-hundreds to low thousands of individuals for both species. Previous studies have shown a correlation between changes in N_e and the population census size over time in birds (e.g. Cosseau et al. 2016, Mueller et al. 2016), suggesting that N_e can be used as a relative measure of population census size; however, the true relationship between N_e and census population size remains unclear and may vary non-linearly (Luikart et al. 2010). In addition, our negative Tajima's D tests for both species suggest that the population sizes of both species could be in recovery (i.e, expansion stage) from a severe or long-lasting bottleneck (Fay and Wu 1999). Given that the population of *Z. murphyi* is endemic to Kolombangara and likely has a small population size, it is imperative to keep the island of Kolombangara in its current state. Kolombangara is one of the most biodiverse islands in the Solomon Archipelago, with two endemic bird species: *Z. murphyi*, and the Kolombangara/Sombre Leaf Warbler *Seicercus amoneus* found at higher elevations (Mayr and Diamond 2001, Weeks et al. 2017). Currently, elevations of 400 meters and above on Kolombangara are protected from logging; however, existing lowland logging

likely has already had a significant effect on lowland species, the microclimate, and species that move between high and low elevations (Weeks et al. 2017). Even though the endemic *Z. murphyi* lives in currently protected habitat, this species is not immune to lowland habitat changes and care should be taken to protect the current population size.

In addition, it is quite concerning that our effective population size estimates of *Z. kulambangrae* based on our Kolombangara genetic samples are so low, given this species is also found on the large neighboring island of New Georgia (2,307 km^2 in size) and the smaller nearby islands of Vangunu (509 km^2), Nggatokae (98km^2), Kohingo/Arundel (~25 km^2), and Vonavona (21km^2) and their satellites (Dutson 2011). Although we lacked genetic samples of *Z. kulambangrae* from other nearby islands within its range, if the individuals from Kolombangara are genetically distinct from other islands, care should be taken to conserve this small, genetically distinct population on Kolombangara. Conversely, if there is gene flow across all these islands within its range, an estimated effective population size of 650-2500 individuals is extremely concerning, given that this estimated effective population size would be representative of the entire species across its full range over a wide geographic area (> 2500 km^2) and would likely imply a fairly small species census population size. In either case, it is extremely important to carefully conserve the remaining forested habitat of *Z. kulambangrae* individuals on Kolombangara and on the neighboring New Georgia islands.

Within the Solomon Islands, destruction of the native tropical forest by logging is a constant threat (Katovi et al. 2012, 2015, 2016). Particularly on Kolombangara, greater than 90% of the habitat below 400 meters has been previously logged or turned into plantation forest (Katovi et al. 2012). These logged forests take decades to recover and do

not fully recover even 50 years after logging (Katovi et al. 2016). *Z. kulambangrae* is particularly susceptible to the effects of logging, given the species is found below 1000 meters and the highest density of individuals seems to be between 200-400 meters (S. Cowles pers obs). However, this species can be found in patches of secondary regrowth forest or in small habitat corridors, suggesting the ability to re-establish once the forest canopy returns in previously logged habitat. Unfortunately, the majority of the neighboring large New Georgia island and the smaller islands around it have been logged given their flatter topography (Katovi et al. 2015), which may be a contributing factor in the low overall estimated effective population size numbers of this species over its wide geographic range.

During fieldwork on Kolombangara, we captured individuals of *Z. murphyi* using passive netting (*Z. kulambangrae* was mostly captured by target netting using song playback, see Cowles and Uy 2019 for more details). Therefore, our recapture rate for *Z. murphyi* individuals can be used as a rough estimate of census population size (assuming *Z. murphyi* individuals can freely move around the island). Out of a total of 59 birds captured in 2016-2017 (we did not obtain a blood sample from one individual, and only had 58 genetic samples), we passively re-captured four of the same birds on different days (all four birds) and in different netting locations (three of four birds), which is a recapture rate of 6.78%. Translating this rate into a census population estimate gives us approximately 870 individuals, which falls in line with our effective population size estimates based on genomic methods.

Finally, our two methods produced different estimates for effective population size. Stairway Plot gave a much higher estimate of effective population size for both

species as a coalescent-based tree method, whereas N_eEstimator gave a smaller estimate of effective population size using the linkage disequilibrium-based method. Both methods may deal with and be affected differently by GBS datasets with missing data. For example, the site frequency spectrum (SFS) in Stairway Plot is likely to be slightly skewed to the left, as filtering steps during sequence and data processing are likely to filter out SNPs that are only found in a few individuals of the sample. Additionally, each method does make certain assumptions about the two species, such as the mating system, mutation rate, and constant population size, which may be untrue. We know that the estimate of N_e can be affected by various factors affecting a real-life non-idealized population, such as gene flow from outside populations, unequal mating success, selection on traits, and underlying population structure (Fisher 1930, Wright 1931, Crow and Kimura 1970, Wang et al. 2016). Therefore, our N_e estimates can be used as a potential indicator of the actual census population size, but we must be cautious in over-interpretation since we do not know the true relationship between N_e and the actual census population size. However, given that both estimates of N_e for the two species were fairly small, the suitable island habitat is reduced, and that we obtained a small population estimate based on recapture rate, it is very possible that the overall census population size is quite small and similar to our N_e estimates (i.e. low thousands to mid-hundreds).

Relatedness estimates and lack of genetic structuring

The fineRADstructure and kinship coefficient estimate analyses found that *Z. murphyi* individuals had higher overall relatedness compared to *Z. kulambagrae*

individuals, even though both species are estimated to have similar effective population sizes. A broader geographic range could explain the reduced relatedness among *Z. kulambangrae* individuals, given a broader range allows for increased genetic variation across the population (e.g. Frankham 1996). In contrast, the difference in individual relatedness levels of each species could be due to differences in genetic bottlenecks in the original founder populations, given each species likely arose from separate colonization events of Kolombangara (Mayr and Diamond 2001). However, founding flock sizes in *Zosterops* species have been suggested to be large enough to counteract strong bottleneck effects (see Clegg et al. 2002, Estoup and Clegg 2003)

We did not find any genetic structuring by transect (Iriri vs. Ringgi) or by elevation (high vs. low) within each species on Kolombangara. This is in contrast to other White-eye species where genetic clustering is apparent even within scale of 10 to 20 km in distance, such as on Reunion Island (Milá et al. 2010, Habel et. al 2013, Bertrand et al. 2014). However, since Kolombangara Island is only 30 km wide and 688 km^2 in size, with much of the area below 400 meters non-native plantation forest or damaged habitat, there may not be enough area for genetic structuring to occur at finer scales on the island, especially given these two species are restricted to particular elevational ranges.

Similar N_e estimates in other Zosterops species

Given that many of the species in the Zosteropidae family are found on small islands (van Balen 2001, Moyle et al. 2009) or in small-range habitats like sky islands (e.g. Husemann et al. 2015), it could be expected that other species of White-eyes have similar population sizes to the two Kolombangara species. Other researchers have

estimated population sizes in *Zosterops* species. For example, Mulwa et al. (2007) did point counts along 119 line transects for *Z. silvanus* in the highly fragmented Taita Hill forests in South-east Kenya. They estimated that the total population size of this species was around 7,120 individuals over all of the fragments of forest (over 6 km^2 total (Brooks et al. 1998)) that they examined; however, this species does tend to show high movement from fragment to fragment (Lens et al. 2002). On the small 0.17 km^2 Heron Island in Australia, the population of Silvereyes (*Z. lateralis chlorocephalus*) fluctuated between 225-483 birds over 1979-1993, and showed evidence of survival affected by density dependence (McCallum et al. 2000). Husemann et al. (2015) used both single point and temporal samples to estimate N_e, in *Z. poligaster* in the East African sky islands in Kenya using microsatellite data, and found that their estimates varied widely depending on the method used (e.g. estimates of between 325-38,249 individuals for the same population). In general, their estimates were reduced and more variable for single time point samples compared to multiple time point estimates. Single time point samples are often smaller than multiple time point estimates and may underestimate N_e (Barker 2011), but tend to be more accurate with unstructured populations (Holleley et al. 2014) like the Kolombangara White-eyes. Similar to Husemann et al. (2015), we also stress that using multiple estimators for N_e gives a better overall picture of a more accurate population size, as these methods do tend to vary widely in their estimates. Further methods should be developed for estimating N_e with SNP GBS data, as we used the only two methods that we know of that could be applied to our dataset.

It could be expected that population sizes and densities would be similar in the other islands of the New Georgia Province (Tetepare, Rendova, Ranongga, Vella

Lavella), each with their own single White-eye species or subspecies with distinct song and plumage (Diamond 1998, Dutson 2011). Unfortunately, the Gizo White-eye (*Z. luteirostris*) on the small neighboring island of Gizo is an endangered species due to habitat destruction, and not many individuals are thought to remain (within the low hundreds) within the small 34 km^2 piece of intact forest (Birdlife International 2019). More research does need to be undertaken in understanding the population sizes and conservation needs of these *Zosterops* species in the Solomon Islands. Overall, given that this "great speciator" Zosteropidae family of birds specializes in expanding into new niches, often restricted to small islands or restricted habitats, conservation of any forest habitats where these species are found is extremely important. Given that forest-dwelling island birds are at the highest risk of extinction (Johnson and Stattersfield 1990), care must be taken to insure these organisms survive into the human-altered future.

Final Summary

In summary, we found that the effective population sizes of both species of *Zosterops* White-eye birds on Kolombangara number in the mid-hundreds to low thousands. Given that *Z. murphyi* is endemic to the higher elevations of Kolombangara, and that Kolombangara is only one of the very few islands that *Z. kulambangrae* is found on (and that this species is more susceptible to the effects of lowland logging and the creation of plantation forests across its range in the New Georgia islands), it is imperative to keep Kolombangara in its current state and to prevent the addition of increased logging.

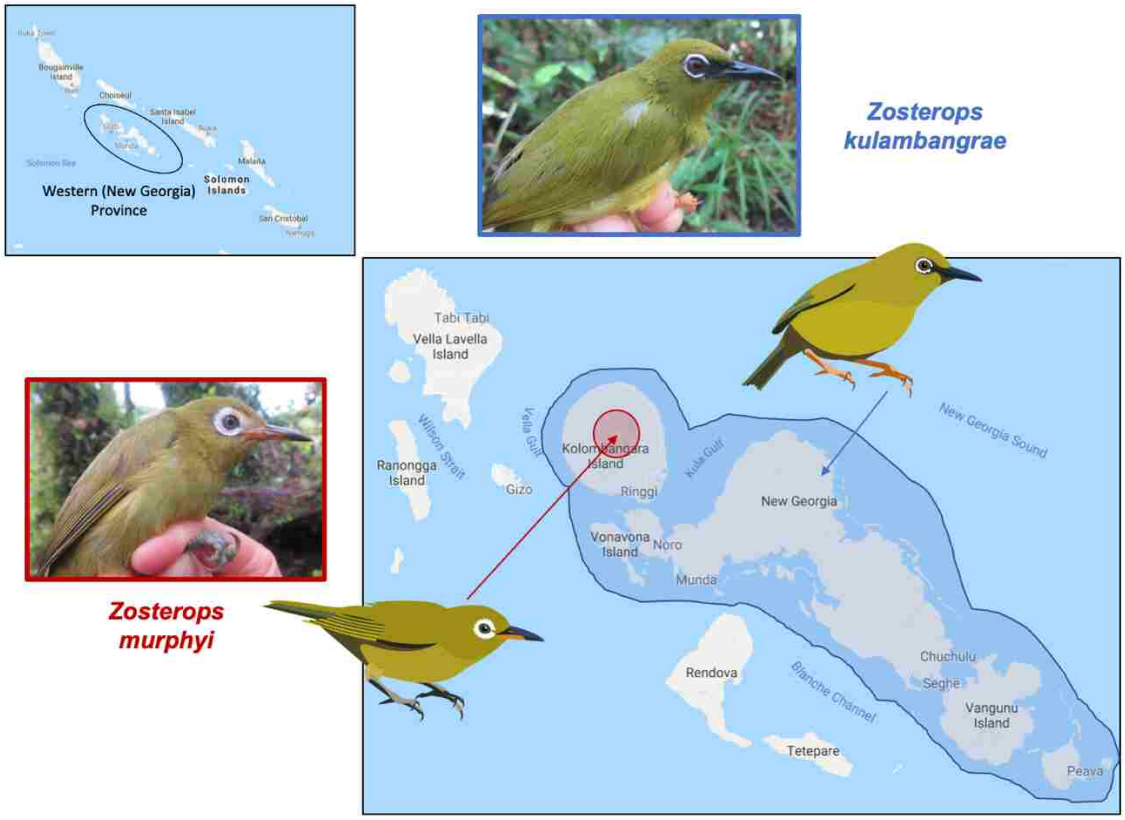

Fig 4.1: The geographic ranges of two species of *Zosterops* White-eye birds found on Kolombangara. *Zosterops kulambangrae* (top, in blue) is found on Kolombangara and across many of the islands in the New Georgia Province. *Z. murphyi* (left, in red) is only found on Kolombangara at elevations greater than 600 meters.

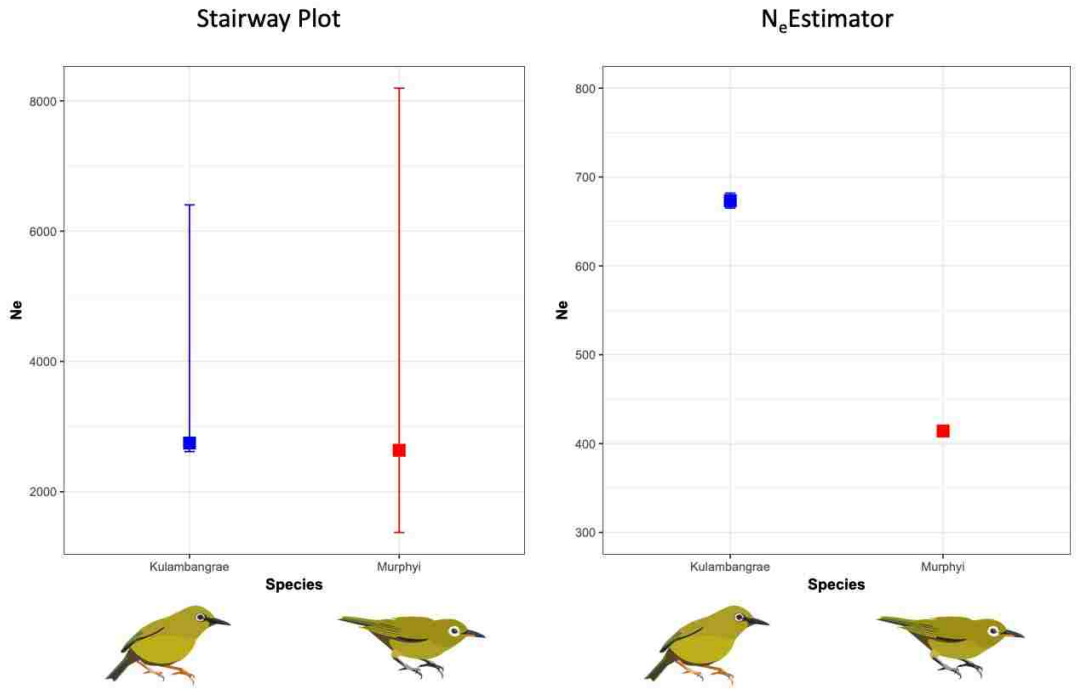

Figure 4.2: Estimated current effective population sizes for *Z. kulambangrae* and *Z. murphyi.* Effective population size for both *Zosterops* species on Kolombangara using the programs Stairway Plot (left) and NeEstimator (right). Error bars represent 95% CI. Please note the change in scaling on the y-axis for each graph.

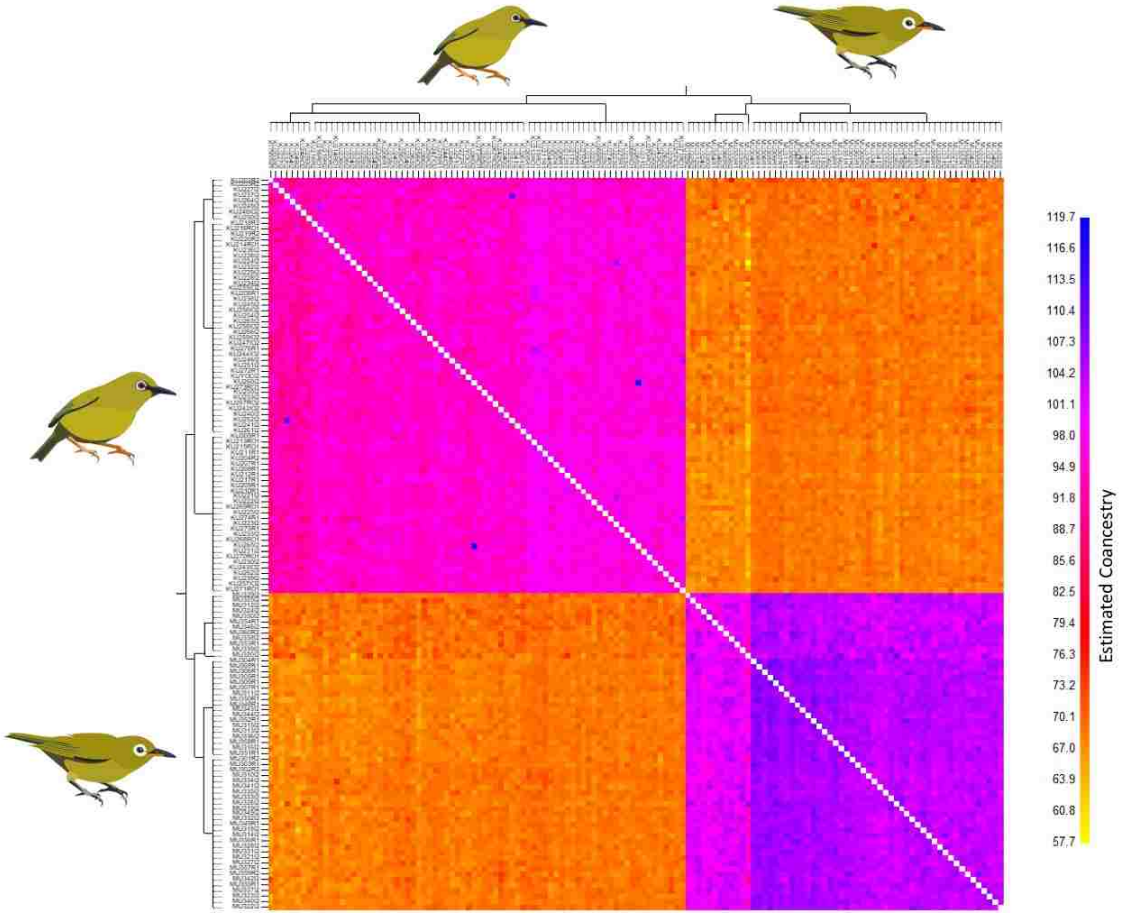

Figure 4.3: fineRADstructure coancestry estimates and population structuring. The fineRADstructure coancestry plot shows *Z. kulambangrae* and *Z. muphyi* are distinct species-specific groups, with *Z. murphyi* individuals showing overall higher coancestry levels compared to *Z. kulambangrae* individuals.

Chapter 5

Summary of Dissertation Findings

The process of speciation can be long and complex, with a multitude of possible trajectories and outcomes (Dobzhansky 1937, Mayr 1942, Coyne & Orr 2004, Price 2008). Studying snapshots of the speciation process in action is valuable in understanding how and why species are formed or go extinct, and the processes responsible. In my Ph.D. dissertation, I focused on the evolutionary dynamics of speciation in two tropical avian systems, applying genomic approaches. I leveraged genomic SNP datasets in two well-known yet understudied clades, the *Amazilia* clade and the *Zosterops* clade, to explore both an early and late stage of the speciation process, respectively.

In Chapter 2, I explored the evolutionary history of the Amazilia Hummingbird. Prior to my study, no studies on species in the *Amazilia* genus had used a genome-wide GBS dataset to infer phylogenetic history, subspecies structuring, or gene flow (e.g. Miller et al, 2011, Ornelas et al. 2013, Rodríguez-Gómez and Ornelas 2018). In my study, I collected 55 blood samples from hummingbirds in Ecuador across six different field sites. Together with 34 tissue samples from hummingbirds across the Peruvian range, I generated a genomic dataset of 34,896 SNPs. From these genomic data, I was able to obtain a clear phylogeny with six distinct subspecies groupings corresponding to the six phenotypically differentiated subspecies. These six subspecies fell into three distinct clades, corresponding with geography: a lowlands Ecuadoran clade (*dumerilii*), a highlands Ecuadorian clade (*alticola* and *azuay*), and finally, a Peruvian clade (*leucophoea*, *amazilia*, and *caeruleigularis*). Using the mitochondrial gene *ND2*, I was able to put an estimated time of divergence at 0.5-3 million years ago for all six

subspecies. These times align with the expected range for divergence times at the subspecies level (Price 2008).

Using additional genomics programs designed to estimate population structuring and gene flow, I found that the three subspecies *dumerilii*, *alticola*, and *leucophea* are genetically similar and have likely had high levels of gene flow. Given these three subspecies have broad geographic distributions and are assumed to come into contact near the Ecuadorian-Peruvian border (Schulenberg et al. 2010, Weller et al. 2019), yet have distinct environmental characteristics (temperature and precipitation) across most of their ranges, I suggest that range expansions into new habitats played an important role in the diversification of these three *A. amazilia* subspecies. In contrast, the subspecies *azuay*, *amazilia*, and *caeruleigularis* are found in habitats that overlap in environmental characteristics with other subspecies, and have smaller range sizes corresponding with geographic features, suggesting geographic isolation instead led to the diversification of these three subspecies. Overall, I was able to demonstrate that multiple evolutionary processes have been working in concert to shape the present-day phylogeny and distribution of the six *A. amazilia* subspecies.

In addition, I found that the subspecies *azuay* (described in Krabbe and Ridgely 2010) is the most genetically distinct and geographically isolated from other subspecies, and shows little to no gene flow with any other subspecies. This means *azuay* could be elevated to species status, and is in contrast to Weller (2000), who suggested that *alticola* should be considered its own separate species (the Loja Hummingbird). As large-scale genomic datasets become more and more common, it is likely that many taxonomic revisions will be necessary, especially at the subspecies and species level. With

increasing levels of genomic resolution, blurring of species and subspecies boundaries is bound to happen, as speciation, especially with some gene flow, leads to complex patterns and relationships among taxa.

In Chapter 3, I examined secondary contact between two closely-related species of *Zosterops* White-eye birds on Kolombangara in the Solomon Islands. My study was the first to examine a contact zone in the rapidly-speciating yet widespread "great speciator" Zosteropidae lineage using a large genomic-wide dataset. In addition, this study was the first in-depth study of an older contact zone greater than 1 million years in this lineage, as *Z. kulambangrae* and *Z. murphyi* are estimated to be just over two million years apart. I found that even with only two million years of separation, there was absolutely no gene flow detected between the two species. In order to put my results of zero hybridization into context with other avian secondary contact zones, I conducted a literature review to find examples of secondary contact in other bird lineages in which a timing of divergence had been estimated, as well as a measure of hybridization. I found that this example was the youngest reported example using genomic data of zero gene flow in birds with overlapping breeding ranges.

My results are even more unusual when put into further context. Within avian secondary contact zones, hybridization tends to be quite common, as to date over 16% of all bird species in the wild have been found to hybridize with a congeneric species or subspecies (Ottenburghs et al. 2015). In addition, it takes a long time to evolve intrinsic reproductive incompatibilities in birds—approximately 5 million years for hybrid sterility, and around 10 million years for hybrid inviability to evolve (Price and Bouvier 2002). Finally, in other secondary contact zones of *Zosterops* species less than 1 million

years old (Mila et al. 2010, Oatley 2012, 2017) and in other species of birds separated for up to 6-7 million years, such as in the Sooty and Cardinal Myzomela (Sardell and Uy 2016) and the Barred Owl and Northern Spotted Owl (Haig et al. 2004), hybridization has been detected. Given these numbers, it would be expected that two *Zosterops* species that are only two million years apart should be hybridizing at least a little, especially given they are found socializing together in mixed flocks within the contact zone—but they are not.

Although it is unknown exactly what reproductive barriers are responsible for keeping the two Kolombangara White-eye species from interbreeding, my study explored some possible mechanisms. From my research, we do know that song and calls are quite distinct between the two species. Similarly, we also know that although overall body plumage color is similar, eye-ring size is quite different between these two species. Plumage and song are often used as pre-mating barriers in birds (e.g. Price 2008, Uy et al. 2009, Uy and Safran 2013, Uy et al. 2018). Finally, physiological adaptions to the habitat or rapid genomic evolution in the *Zosterops* lineage (see Cornetti et al. 2015, Leroy et al. 2019) could also play a role in reproductive isolation. Further genomic studies of rapidly speciating yet widespread lineages, particularly in older secondary contact zones in Zosteropidae, will lend insight into why this lineage is unique in its ability to quickly evolve reproductive isolation yet spread over a wide geographic area.

In Chapter 4, I used my genomic SNP dataset from the Kolombangara *Zosterops* White-eyes to estimate the current population size of each species using two genomic methods designed to work with single time point sampling population data. This study has far-ranging implications for conservation in that region, given that over 90% of the

island habitat has been affected in some way by logging (Katovi et al. 2012). Most noteworthy, I show that the current effective population size of *Z. kulambangrae* is estimated to be similar to the endemic, range-restricted *Z. murphyi*, even though *Z. kulambangrae* is found over a much wide geographic range across most of the New Georgia Islands. However, given that *Z. kulambangrae* is found at lower elevations that are prime targets for easily logging habitat, it is likely that population numbers across all islands within its range have been affected by logging. Given other geographically-isolated White-eye species are on neighboring islands (e.g. like the endangered Gizo White-eye *Z. splendidus* (Dutson 2011)), and that many species within the Zosteropidae family as a whole are native to small islands or restricted ranges (van Balen 2001, Moyle et al. 2009), my study shows that small genomic datasets could possibly be used to estimate population sizes of species in this group, especially when it is not feasible to do long-term field monitoring population studies. Monitoring these population sizes is especially important for conserving biodiversity since small island populations of birds have the highest risk of going extinct (Johnson and Stattersfield 1990).

In addition, my results highlight the need for knowledge about local and endemic species when planning and developing for the future. For example, there is constant communication among locals on Kolombangara about logging, which is extremely lucrative in the short-term but is extremely destructive with impacts lasting for decades (Katovi et al. 2016), versus preserving the forest and finding alternative ways to earn income. At least on Kolombangara, many local people are aware of the negative impacts previous logging has had on the island, and there is a drive to push for the creation of a national park on land above 400 meters in elevation to preserve endemic species on the

island (J. Ghemu pers. comm.). In addition, other ways of earning income, like a water bottling factory (J. Ghemu pers. comm.) are being explored on the island as potential ways to take advantage of the natural resources while reducing the risk of negative impacts on wildlife. Further study of endemic and limited range species must be done to inform locals about conservation and to help guide them when making decisions about their land.

Across all three data chapters of my dissertation, I used genomic datasets combined with a variety of genomics programs to explore patterns of phylogeography, gene flow, population structuring, and population size. One main conclusion from this process is that it is best to use a variety of genomic programs to fully assess the outcome and give a comprehensive picture of each story. Multiple methods gave slightly contrasting or inconclusive results alone, but overall, more solid conclusions can be drawn from a redundancy of methods. Genomic datasets are an incredibly rich source of information, and large-scale genome-wide datasets can be used to answer a multitude of questions.

In summary, my dissertation exemplifies the importance of studying different stages of speciation, especially across highly mobile organisms such as birds. Using genomic data in concert with phenotypic and environmental datasets is a powerful tool to answer questions about speciation. My chapters are the first studies to use genome-wide SNP datasets in the *Amazilia* and *Zosterops* clades to answer underlying evolutionary questions about speciation, which can inform us about the generation, maintenance, and loss of biodiversity.

REFERENCES

Allen, J.L., and J.C. Lendemer. 2016. Climate change impacts on endemic, high-elevation lichens in a biodiversity hotspot. Biodivers. Conserv. 25:555-568.

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.

Arbogast, B.S., S.V. Drovetski, R.L. Curry, P.T. Boag, G. Seutin, P.R. Grant, B.R. Grant, and D.J. Anderson. 2006. The origin and diversification of Galapagos mockingbirds. Evolution 60:370–382.

Arnold, M.L., S. Tang, S.J. Knapp, and N.H Martin. 2010. Asymmetric introgressive hybridization among Louisiana Iris species. Genes 1:9-22.

Axelsson, E., N.G.C. Smith, H. Sunström, S. Berlin, and H. Ellegren. 2004. Male-biased mutation rate and divergence in autosomal, Z-linked and W-linked introns of chicken and turkey. Mol. Biol. Evol. 21:1538-1547.

Baldassarre, D.T., T.A. White, J. Karubian, and M.S. Webster. 2014. Genomic and morphological analysis of a semipermeable avian hybrid zone suggests asymmetrical introgression of a sexual signal. Evolution 68:2644-2657.

Barker, J.S.F. 2011. Effective population size of natural populations of *Drospohila buzzatii*, with a comparative evaluation of nine methods of estimation. Mol. Ecol. 20:4452-4471.

Barluenga, M., K.N. Stölting, W. Salzburger, M. Muschick, and A. Meyer. 2006. Sympatric speciation in Nicaraguan crater lake cichlid fish. Nature 439:719-723.

Barraclough, T. G., P. H. Harvey, and S. Nee. 1995. Sexual selection and taxonomic diversity in passerine birds. Proc. R. Soc. Lond. B 259:211– 215.

Barrera-Guzmán, A.O., A. Aleixo, M.D. Shawkey, and J.T. Weir. 2017. Hybrid speciation leads to novel male secondary sexual ornamentation of an Amazonian bird. Proc. Natl. Acad. Sci. USA 115:E218–25.

Barton, N.H., and G.M. Hewitt. 1985. Analysis of hybrid zones. Ann. Rev. Ecol. Syst.16:113-148.

Baxter, S.W., J.W. Davey, J.S. Johnston, A.M. Shelton, D.G. Heckel, C.D. Jiggins, and M.L. Blaxter. 2011. Linkage mapping and comparative genomics using next-generation RAD sequencing of a non-model organism. PLoS One 6:e19315.

Beckman, E.J., and C.C. Witt. 2015. Phylogeny and biogeography of the New World siskins and goldfinches: rapid, recent diversification in the Central Andes. Mol. Phylogenet. Evol. 87:28-45.

Benning, T.L., D. LaPointe, C.T. Atkinson, and P.M. Vitousek. 2002. Interactions of climate change with biological invasions and land use in the Hawaiian Islands: modeling the fate of endemic birds using a geographic information system. PNAS 99:14246-14249.

Bentley, D.R. 2006. Whole-genome re-sequencing. Curr. Opin. Genetics Dev. 16:545-552.

Bertrand, J.A.M., Y.X.C. Bourgeois, B. Delahaie, T. Duval, R. García-Jiménez, J. Cornuault, P. Heeb, B. Milá, B. Pujol, and C. Thébaud. 2014. Extremely reduced dispersal and gene flow in an island bird. Heredity 112:190-196.

Bioacoustics Research Program. 2014. Raven Pro: Interactive Sound Analysis Software (Version 1.5) [Computer software]. Ithaca, NY: The Cornell Lab of Ornithology. Available from http://www.birds.cornell.edu/raven.

BirdLife International. 2019. Species factsheet: *Zosterops luteirostris*. Downloaded from http://www.birdlife.org on 13/09/2019.

Bleiweiss, R., J.A. Kirsch, and J.C. Matheus. 1997. DNA hybridization evidence for the principal lineages of hummingbirds (Aves: Trochilidae). Mol. Biol. Evol. 14:325-43.

Block, N.L., S.M. Goodman, S.J. Hackett, J.M. Bates, and M.J. Raherilalao. 2015. Potential merger of ancient lineages in a passerine bird discovered based on evidence from host-specific ectoparasites. Ecol. Evol. 5:3743-3755.

Bradbury, P.J., Z. Zhang, D.E. Kroon, T.M. Casstevens, Y. Ramdoss, and E.S. Buckler. 2007. TASSEL: Software for association mapping of complex traits in diverse samples. Bioinformatics 23:2633-2635.

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

Brelsford, A., B. Milá, and D.E. Irwin. 2011. Hybrid origin of Audubon's warbler. Mol. Ecol. 20:2380-2389.

Brooks, T., L. Lens, J. Barnes, R. Barnes, J.K. Kageche, and C. Wilder. 1998. The conservation status of the forest birds of the Taita Hills, Kenya. Bird Conserv. Int. 8:119-139.

Calviño-Cancela, M. 2006. Time-activity budgets and behavior of the Amazilia Hummingbird *Amazilia amazilia* (Apodiformes: Trochilidae) in an urban environment. Int. J. Trop. Biol. 54:873-878.

Carling, M.D. and R.T. Brumfield. 2008. Haldane's rule in an avian system: using cline theory and divergence population genetics to test for differential introgression of mitochondrial, autosomal, and sex-linked loci across the *Passerina* bunting hybrid zone. Evolution 62:2600-2615.

Caro, S.P., A. Charmantier, M.M. Lambrechts, J. Blondel, J. Balthazart, and T.D. Williams. 2009. Local adaptation of timing of reproduction: females are in the driver's seat. Funct. Ecol. 23:172-179.

Claramunt, S., E.P. Derryberry, J.V. Remsen Jr., and R.T. Brumfield. 2012. High dispersal ability inhibits speciation in a continental radiation of passerine birds. Proc. R. Soc. B 279:1567-1574.

Clegg, S.M., S.M. Degnan, J. Kikkawa, C. Moritz, A. Estoup, and I.P.F. Owens. 2002. Genetic consequences of sequential founder effects in an island bird. Proc. Natl. Acad. Sci. USA 99:8127-8132

Clegg, S.M. and A.B. Phillmore. 2010. The influence of gene flow and drift on genetic and phenotypic divergence in two species of Zosterops on Vanuatu. Phil. Trans. R. Soc. B 365:1077-1092.

Cooney, C.R., J.A. Tobias, J.T. Weir, C.A. Botero, and N. Seddon. 2017. Sexual selection, speciation and constraints on geographical range overlap in birds. Ecol. Lett. 20:863-871.

Cornetti, L., L.M. Valente, L.T. Dunning, X. Quan, R.A. Black, O. Hébert, and V. Savolainen. 2015. The genome of the "great speciator" provides insights into bird diversification. Genome Biol. Evol. 7:2680-2691.

Cornuault, J., B. Delahaie, J.A.M. Bertrand, Y.X.C. Bourgeois, B. Milá, P. Heeb, and C. Thébaud. 2015. Morphological and plumage colour variation in the Réunion grey white-eye (Aves: *Zosterops borbonicus*): assessing the role of selection. Biol. J. Linnean Soc. 114:459-473.

Cosseau, L., M. Husemann, R. Foppen, C. Vangestel, and L. Lens. 2016. A longitudinal genetic survey identifies temporal shifts in the population structure of Dutch house sparrows. Heredity 117:259-267.

Cowles, S.A., and J.A.C. Uy. 2019. Rapid, complete reproductive isolation in two closely-related White-eye bird species despite broadly overlapping ranges. Evolution 73:1647-1662.

Cox, S.C., R.P. Prys-Jones, J.C. Habel, B.A. Amakobe, and J.J. Day. 2014. Niche divergence promotes rapid diversification of East African sky island white-eyes (Aves: Zosteropidae). Mol. Ecol. 23:4103-4118.

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

Coyne, J.A., and T.D. Price. 2000. Little evidence for sympatric speciation in island birds. Evolution 54:2166-2171.

Crow, J. F. and M. Kimura. 1970. An introduction to population genetics theory. New York: Harper and Row.

Darwin, C. 1859. On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life. W. Clowes & Sons: London.

Diamond J. 1998. Geographic variation in vocalizations of the White-eye superspecies *Zosterops [griseotinctus]* in the New Georgia group. Emu 98:70-74.

Diamond J.M., M.E. Gilpin, and E. Mayr. 1976. Species-distance relation for birds of the Solomon Archipelago, and the paradox of the great speciators. Proc. Natl. Acad. Sci. USA 73:2160-2164.

Dieckmann, U., and M. Doebeli. 1999. On the origin of species by sympatric speciation. Nature 400:354-357.

Dirnböck, T., F. Essl, and W. Rabitsch. 2011. Disproportional risk for habitat loss of high-altitude endemic species under climate change. Glob. Change Biol. 17:990-996

Dirzo, R., H.S. Young, M. Galetti, G. Ceballos, N.J.B. Issac, and B. Collen. 2014. Defaunation in the Anthropocene. Science 345:401-406.

Do, C., R.S. Waples, D. Peel, G.M. Macbeth, B.J. Tillett, and J. R. Ovenden. 2014. NeEstimator V2: re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. Mol. Ecol. Resour. 14:209-214.

Dobzhansky, T. 1935. A critique of the species concept in biology. Philos. Sci. 2:344-355.

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

Durand, E.Y., N. Patterson, D. Reich, and M. Slatkin. 2011. Testing for ancient admixture between closely related populations. Mol. Biol. Evol. 28:2239–2252.

Dutson, G. 2011. Birds of Melanesia: Bismarcks, Solomons, Vanuatu, and New Caledonia. Bloomsbury Publishing, London.

Elias, M., M. Joron, K. Willmott, K.L. Silva-Brandão, V. Kaiser, C.F. Arias, L.M. Gomez Piñerez, S. Uribe, A.V.Z. Brower, A.V.L. Frietas, and C.D. Jiggins. 2009. Out of the Andes: patterns of diversification in clearwing butterflies. Mol. Ecol. 18:1716-1729.

Elshire, R.J., J.C. Glaubitz, Q. Sun, J.A. Poland, K. Kawamoto, E.S. Buckler, and S.E. Mitchell. 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS One 6:e19379.

Endler, J.A. 1990. On the measurement and classification of colour in studies of animal colour patterns. Biol. J. Linnean Soc. 41:315-352.

Estoup, A., and S.M. Clegg. 2003. Bayesian inference on the recent island colonization history by the bird *Zosterops lateralis lateralis*. Mol. Ecol. 12:657-674.

Falla, R.A., R.B. Sibson, and E.G. Turbott. 1966. A field guide to the birds of New Zealand and outlying islands. Collins, London.

Fay, J.C., and C.-I Wu. 1999. A human population bottleneck can account for the discordance between patterns of mitochondrial versus nuclear DNA variation. Mol. Biol. Evol. 17:1003-1005.

Fick, S.E., and R.J. Hijmans. 2017. Worldclim 2: New 1-km spatial resolution climate surfaces for global land areas. Int. J. Climatol. 37:4302-4315.

Fisher, R.A. 1930. The genetical theory of natural selection. Oxford University Press.

Francis, R. M. 2016. POPHELPER: An R package and web app to analyse and visualise population structure. Molecular Ecology Resources. DOI: 10.1111/1755-0998.12509

Frankham, R. 1996. Relationship of genetic variation to population size in wildlife. Conserv. Biol. 10:1500-1508.

Freeman, B.G. 2015. Competitive interactions upon secondary contact drive elevational divergence in tropical birds. Am. Nat. 186:470-479.

Funk, D.J., and K.E. Omland. 2003. Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. Annu. Rev. Ecol. Evol. Syst. 34:397-423.

Gaston, K.J. 2000. Global patterns in biodiversity. Nature 405:220-227.

Gavrilets, S., and J. B. Losos. 2009. Adaptive radiation: contrasting theory with data. Science 323:732-737.

Genovart, M., D. Oro, J. Juste, and G. Bertorelle. 2007. What genetics tell us about the conservation of the critically endangered Balearic Shearwater. Biol. Cons. 137:283-293.

Gill, F.B. 1970. Hybridization in Norfolk white-eyes (*Zosterops*). Condor 72:481-482.

Gill, F.B. 1973. Intra-island variation in the Mascarene White-eye *Zosterops borbonica*. Ornithol. Monogr. 12:1-66.

Glaubitz, J.C., T.M. Casstevens, F. Lu, J. Harriman, R.J. Elshire, Q. Sun, and E.S. Buckler. 2014. TASSEL-GBS: A high capacity genotyping by sequencing analysis pipeline. PLoS One 9: e90346.

Glotzbecker, G.J., D.M. Walters, and M.J. Blum. 2016. Rapid movement and instability of an invasive hybrid swarm. Evol. Appl. 9:741-755.

Goudet, J. 2005. HIERFSTAT, a package for R to compute and test hierarchical F-statistics. Mol. Ecol. Notes 5:184-186.

Goudet J., and T. Jombart. 2015. Hierfstat: Estimation and tests of hierarchical F-statistics. R package version 0.04-22. Retrieved from https://CRAN.R-project.org/package=hierfstat

Graham, C.H., J.L. Parra, C. Rahbek, and J.A. McGuire. 2009. Phylogenetic structure in tropical hummingbird communities. PNAS 106:19673-19678.

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

Grant, B.R., and P.R. Grant. 1996. Cultural inheritance of song and its role in the evolution of Darwin's Finches. Evolution 50:2471-2487.

Green, R.E., J. Krause, A.W. Briggs, T. Maricic, U. Stenzel, M. Kircher, N. Patterson, H. Li, W. Zhai, M.H. Fritz, et al. 2010. A draft sequence of the Neandertal genome. Science 328:710–722.

Habel, J.C., S. Cox, F. Gassert, R.K. Mulwa, J. Meyer, and L. Lens. 2013. Population genetics of the East African white-eye species complex. Conserv. Genet. 5:1019-1028.

Habel, J.C., M. Husemann, and W. Ulrich. 2015. Patterns of contact call differentiation in the panmictic East African Abyssinian white-eye *Zosterops abyssinicus* (Aves: Passeriformes). Ecol. Evol. 5:5974-5982.

Haig, S.M., T.D. Mullins, E.D. Forsman, P.W. Trail, and L. Wennerberg. 2014. Genetic identification of Spotted owls, Barred owls, and their hybrids: legal implications of hybrid identity. Conserv. Biol. 18:1347-1357.

Hardy, O.J., and X. Vekemans. 2002. SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. Mol. Ecol. Notes 2:618-620.

Hansson, B., M. Tarka, D.A. Dawson, and G. J. Horsburgh. 2012. Hybridization but no evidence for backcrossing and introgression in a sympatric population of Great Reed Warblers and Clamorous Reed Warblers. PLoS ONE 7:e31667.

Harvey, M.G., and Brumfield, R.T. 2015. Genomic variation in a widespread Neotropical bird (*Xenops minutus*) reveals divergence, population expansion, and gene flow. Mol. Phylogenet. Evol. 83:305-316.

Helm-Bychowski, K.M., and A.C. Wilson. 1986. Rates of nuclear DNA evolution in pheasant-like birds: evidence from restriction maps. Proc. Natl. Acad. Sci. USA 83:688-692.

Hermansen, J.S., F. Haas, C.N. Trier, R.I. Bailey, A.J. Nederbragt, A. Marzal, and G.-P. Sætre. 2014. Hybrid speciation through sorting of parental incompatibilities in Italian sparrows. Mol. Ecol. 23:5831-5842.

Hijmans, R.J., and J. van Etten. 2014. raster: Geographic data analysis and modeling (R package). https://CRAN.R-project.org/package=raster

Hill, G.E. 2017. The mitonuclear compatibility species concept. Auk 134:393-409.

Hogner, S., T. Laskemoen, J.T. Lifjeld, J. Porkert, O. Kleven, T. Albayrak, B. Kabasakal, and A. Johnsen. 2012. Deep sympatric mitochondrial divergence without reproductive isolation in the common redstart *Phoenicurus phoenicurus*. Ecol. Evol. 2:2974-2988.

Holleley, C.E., R.A. Nichols, M.R. Whitehead, A.T. Adamack, M.R. Gunn, and W.B. Sherwin. 2014. Testing single-sample estimators of effective population size in genetically structured populations. Conserv. Genet. 15:23-35.

Huang, J-P. 2019. Holocene population decline and conservation implication for the Western Hercules Beetle, *Dynastes grantii*. J. Hered. 110:629-637.

Hudson, E.J., and T.D. Price. 2014. Pervasive reinforcement and the role of sexual selection in speciation. J. Hered. 105:821-833.

Hugall, A.F., and D. Stuart-Fox. 2012. Accelerated speciation in colour-polymorphic birds. Nature 485:631-634.

Hughes, C., and R. Eastwood. 2006. Island radiation on a continental scale: exceptional rates of plant diversification after uplift of the Andes. PNAS 103:10334-10339.

Husemann, M., W. Ulrich, and J.C. Habel. 2014. The evolution of contact calls in isolated and overlapping populations of two white-eye congeners in East Africa (Aves, *Zosterops)*. BMC Evol. Biol. 14:115.

Hutter, C.R., J.M. Guayasamin, and J. J. Wiens. 2013. Explaining Andean megadiversity: the evolutionary and ecological causes of glassfrog elevational richness patterns. Ecol. Lett. 16:1135-1144.

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

Jetz, W., G. H. Thomas, J.B. Joy, K. Hartmann, and A. O. Mooers. 2012. The global diversity of birds in space and time. Nature 491:444-448.

Jiménez, R.A., and J.F. Ornelas. 2016. Historical and current introgression in a Mesoamerican hummingbird species complex: a biogeographic perspective. PeerJ 4:e1556.

Jombart, T., and I. Ahmed. 2011. adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. Bioinformatics 27:3017-3071.

Johnson, C.N., A. Balmford, B.W. Brook, J.C. Buettel, M. Galetti, L. Guangchun, and J.M. Wilmshurst. 2017. Biodiversity losses and conservation responses in the Anthropocene. Science 356:270-275.

Johnson, K.P., and J. Seger. 2001. Elevated rates of nonsynonymous substitution in islands birds. Mol. Biol. Evol. 18:874-881.

Johnson, T.H. and A.J. Stattersfield. 1990. A global review of island endemic birds. Ibis 132:167-180.

Katch, E.K., E.M. Kierepka, J.R. Heffelfinger, and O.E. Rhodes Jr. 2011. Hybrid swarm between divergent lineages of mule deer (*Odocoileus hemionus*). Mol. Ecol. 20:5265-5279.

Katovai, E., A.L. Burley, and M.M. Mayfield. 2012. Understory plant species and functional diversity in the degraded wet tropical forests of Kolombangara Island, Solomon Islands. Biol. Cons. 145:214–224.

Katovi, E., W. Edwards, and W.F. Laurance. 2015. Dynamics of logging in the Solomon Islands: the need for restoration and conservation alternatives. Trop. Conserv. Sci. 8:718-731.

Katovai, E., M. Sirikolo, U. Srinivasan, W. Edwards, and W.F. Laurance. 2016. Factors influencing tree diversity and compositional change across logged forests in the Solomon Islands. For. Ecol. Manag. 372:53-63.

Keinan, A., J.C. Mullikin, N. Patterson, and D. Reich. 2007. Measurement of the human allele frequency spectrum demonstrates greater genetic drift in East Asians than in Europeans. Nat. Genet. 39:1251–1255.

Knaus, B.J., and N.J. Grunwald. 2017. VCFR: a package to manipulate and visualize variant call format data in R. Mol. Ecol. Resour. 17:44-53.

Koboldt, D.C., K.M. Steinberg, D.E. Larson, R.K. Wilson, and E.R. Maradis. 2013. The next-generation sequencing revolution and its impact on genomics. Cell 155:27-38.

Körner, C., and E. M. Spehn. 2002. Mountain biodiversity: a global assessment. Taylor and Francis: London, UK.

Krabbe, N., and R.S. Ridgely. 2010. A new subspecies of Amazilia Hummingbird *Amazilia amazilia* from southern Ecuador. Bull. Br. Orn. Club. 130:3-7.

Kumar, S., G. Stecher, and K. Tamura. 2016. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger data sets. Mol. Biol. Evol. 33:1870-1874

Lanfear, R., S.Y.W. Ho, D. Love, and L. Bromham. 2010. Mutation rate is linked to diversification in birds. Proc. Natl. Acad. Sci. USA 107:20423-20428.

Lavretsky, P., A. Engilis Jr., J.M. Eadie, and J.L. Peters. 2015. Genetic admixture supports an ancient hybrid origin of the endangered Hawaiian duck. J. Evol. Biol. 28:1005-1015.

Lawson, D.J., G. Hellenthal, S, Myers, and D. Falush. 2012. Inference of population structure using dense haplotype data. PLoS Genet. 8:e1002453.

Lawson, L.P., B. Fessl, F.H. Vargas, H.L. Farrington, H.F. Cunninghame, J.C. Mueller, E. Nemeth, P.C. Sevilla, and K. Petren. 2017. Slow motion extinction: inbreeding, introgression, and loss in the critically endangered mangrove finch (*Camarhynchus heliobates*). Conserv. Genet. 18:159-170.

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

Lens, L., S. Von Dongen, K. Norris, M. Githiru, and E. Matthysen. 2002. Avian persistence in fragmented rainforest. Science 259:1236-1238.

Leroy, T., Y. Anselmetti, M-K. Tilak, S. Bérard, L. Csukonyi, M. Gabrielli, C. Scornavacca, B. Milá, C. Thébaud, and B. Nabholz. 2019. A bird's white-eye view on neosex chromosome evolution. bioRxiv:https://doi.org/10.1101/505610.

Li, H., and R. Durbin. 2011. Inference of human population history from individual whole-genome sequences. Nature 475:493-496.

Liou, L.W., and T.D. Price. 1994. Speciation by reinforcement of premating isolation. Evolution 48:1451-1459.

Lischer, H.E.L., and L. Excoffier. 2011. PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. Bioinformatics 28:298-299.

Liu S., A.L. Ferchaud, P. Groenkjaer, R. Nygaard, and M.M. Hansen. 2018. Genomic parallelism and lack thereof in contrasting systems of three-spine sticklebacks. Mol. Ecol. 27:1-19.

Liu X. and Y.X. Fu. 2015. Exploring population size changes using SNP frequency spectra. Nat. Genet. 47:555-559.

Loarie, S.R., B.E. Carter, K. Hayhoe, S. McMahon, R. Moe, C.A. Knight, and D.D. Ackerly. 2008. Climate change and the future of California's endemic flora. PLoS ONE 3:e2502.

Loiselle, B.A., V.L. Sork, J. Nason, and C. Graham. 1995. Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). Am. J. Bot. 82:1420-1425.

Longmire, J. L., M. Maltbie, and R.J. Baker. 1997. Use of" lysis buffer" in DNA isolation and its implication for museum collections. Museum of Texas Tech University.

Luikart, G., N. Ryman, D.A. Tallmon, M.K. Schwartz, and F.W. Allendorf. 2010. Estimation of census and effective population sizes: the increasing usefulness of DNA-based approaches. Conserv. Genet. 11:355-373.

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.

Malinsky, M., E. Trucchi, D.J. Lawson, and D. Falush. 2018. RADpainter and fineRADstructure: Population inference from RADseq Data. Mol. Biol. Evol. 35:1284-1290.

Malinsky. M. 2019. DSuite--fast D-statistics and related admixture evidence from vcf files. BioRxiv Preprint. Available from doi**:** https://doi.org/10.1101/634477

Mamanova, L., A.J. Coffey, C.E. Scott, I. Kozarewa, E.H. Turner, A. Kumar, E. Howard, J. Shendure, and D.J. Turner. 2010. Target-enrichment strategies for next-generation sequencing. Nat. Methods 7:111-118.

Mardis, E.R. 2008. The impact of next-generation sequencing technology on genetics. Trends Genet. 24:133-141.

Mayr, E. 1942. Systematics and the origin of species from the viewpoint of a zoologist. Harvard University Press, Cambridge, Massachusetts

Mayr, E., and J. Diamond. 2001. The birds of northern Melanesia. Oxford University Press, New York, NY.

McCallum, H., J. Kikkawa, and C. Catterall. 2000. Density dependence in an island population of silvereyes. Ecol. Lett. 3:95-100.

McGuire, J.A., C.C. Witt, J.V. Remsen, Jr., A. Corl, D.L. Rabosky, D.L. Altshuler, and R. Dudley. 2014. Molecular phylogenetics and the diversification of hummingbirds. Curr. Biol. 24:1-7.

McKay, B.E., and R.M. Zink. 2010. The causes of mitochondrial DNA gene tree paraphyly in birds. Mol. Phylogenet. Evol. 54:647-650.

Mees, G.F. 1969. A systematic review of the Indo-Australian Zosteropidae (Part III). Zool. Verh. 102:1-390.

Melo, M., B.H. Warren, and P.J. Jones. 2011. Rapid parallel evolution of aberrant traits in the diversification of the Gulf of Guinea white-eyes (Aves, Zosteropidae). Mol. Ecol. 20:4953-4967.

Milá, B., B.H. Warren, P. Heeb, and C. Thébaud. 2010. The geographic scale of diversification on islands: genetic and morphological divergence at a very small spatial scale in the Mascarene grey white-eye (Aves: *Zosterops borbonicus*). BMC Evol. Biol. 10:158.

Miller, M.J., M.J. Lelevier, E. Bermingham, J.T. Klicka, P.E. Escalante, and K. Winkler. 2011. Phylogeography of the Rufous-tailed hummingbird (*Amazilia tzactl*). Condor 113:806-816.

Money, D., K. Gardner, Z. Migicovsky, H. Schwaninger, G-Y. Zhong, and S. Myles. 2015. LinkImpute: fast and accurate genotype imputation for nonmodel organisms. G3-Genes Genom. Genet. 5:2383-2390.

Moore, W.S., and D.B. Buchanan. 1985. Stability of the Northern flicker hybrid zone in historical times: implications for adaptive speciation theory. Evolution 39:135-151.

Moore, I.T., F. Bonier, and J.C. Wingfield. 2005. Reproductive asynchrony and population divergence between two tropical bird populations. Behav. Ecol. 16:755-762.

Morales, H.E., A. Pavlova, L. Joseph, and P. Sunnucks. 2015. Positive and purifying selection in mitochondrial genomes of a bird with mitonuclear discordance. Mol. Ecol. 24:2820-2837.

Moyle R.G., C.E. Filardi, C.E. Smith, and J. Diamond. 2009. Explosive Pleistocene diversification and hemispheric expansion of a “great speciator.” Proc. Natl. Acad. Sci. USA 106:1863-1868.

Mueller, A-K., N. Chakarov, O. Kruger, and J.I. Hoffman. 2016. Long-term effective population size dynamics of an intensively monitored vertebrate population. Heredity 117:290-299.

Mulwa, R.K., L.A. Bennun, C.K.P.O. Ogol, and L. Lens. 2007. Population status and distribution of Taita White-eye Zosterops silvanus in the fragmented forests of Taita Hills and Mount Kasigau, Kenya. Bird Conserv. Int. 17:141-150.

Nabholz, B., A. Künster, R. Wang, E.D. Jarvis, and H. Ellegren. 2011. Dynamic evolution of base composition: causes and consequences in avian phylogenomics. Mol. Biol. Evol. 28:2197-2210.

Nadachowska-Brzyska, K., R. Burri, L. Smeds, and H. Ellegren. 2016. PSMC analysis of effective population sizes in molecular ecology and its application to black-and-white *Ficedula* flycatchers. Mol. Ecol. 25:1058-1072.

Nam, K., C. Mugal, B. Nabholz, H. Schielzeth, J.B.W. Wolf, N. Backström, A. Künstner, C.N. Balakrishnan, A. Heger, C.P. Ponting, D.F. Clayton, and H. Ellegren. 2010. Molecular evolution of genes in avian genomes. Genome Biol. 11:R68.

North, A.J. 1904. Ornithological Notes. Rec. Aust. Mus. 5:337-338.

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

Nunziata, S.O., and D.W. Weisrock. 2018. Estimation of contemporary effective population size and population declines using RAD sequence data. Heredity 120:196–207.

Oatley, G., G. Voelker, T.M. Crowe, and R.C.K. Bowie. 2012. A multi-locus phylogeny reveals a complex pattern of diversification related to climate and habitat heterogeneity in southern African white-eyes. Mol. Phylogenetics Evol. 64:633-644.

Oatley, G., D.H. De Swardt, R.J. Nuttall, T.M. Crowe, and R.C.K. Bowie. 2017. Phenotypic and genotypic variation across a stable white-eye (*Zosterops* sp.) hybrid zone in central South Africa. Biol. J. Linn. Soc. 121:670-684.

O'Connell, D.P., D.J. Kelly, N. Lawless, K. O'Brien, F. Ó. Marcaigh, A. Karya, K. Analuddin, and N.M. Marples. 2019. A sympatric pair of undescribed white-eye species (Aves: Zosteropidae: Zosterops) with different origins. Zool. J. Linn. Soc. 186:701-724.

Oliveira, R., E. Randi, F. Mattucci, J.D. Kurushima, L.A. Lyons, and P.C. Alves. 2015. Toward a genome-wide approach for detecting hybrids: informative SNPs to detect introgression between domestic cats and European wildcats (*Felis silvestris*). Heredity 115:195-205.

Olson, J.R., S.J. Cooper, D.L. Swanson, M.J. Braun, and J.B. Williams. 2010. The relationship of metabolic performance and distribution in Black-capped and Carolina Chickadees. Physiol. Biochem. Zool. 83:263-275.

Ornelas, J.F., C. González, A. Espinosa de los Monteros, F. Rodríguez-Gómez, and L.M. García-Feria. 2014. In and out of Mesoamerica: temporal divergence of Amazilia hummingbirds pre-dates the orthodox account of the completion of the Isthmus of Panama. J. Biogeogr. 41:168-181.

Oswald, J.A., M.G. Harvey, R.C. Remsen, D.U. Foxworth, S.W. Cardiff, D.L. Dittman, L.C. Megna, M.D. Carling, and R.T. Brumfield. 2016. Willet be one species or two? A genomic view of the evolutionary history of *Tringa semipalmata*. Auk 133:593-614.

Ottenburghs, J., R. C. Ydenberg, P. Van Hooft, S. E. Van Wieren, and H. H. Prins. 2015. The Avian Hybrids Project: gathering the scientific literature on avian hybridization. Ibis 157:892–894.

Paradis, E. 2010. pegas: an R package for population genetics with an integrated-modular approach. Bioinformatics 26:419-420.

Payseur, B.A., and L.H. Rieseberg. 2016. A genomic perspective on hybridization and speciation. Mol. Ecol. 25:2337-2360.

Pérez-Escobar, O.A., G. Chomicki, F. Condamine, A.P. Karremans, D. Bogarín, N.J. Matzke, D. Silvestro, and A. Antonelli. 2017. Recent origin and rapid speciation of Neotropical orchids in the world's richest plant biodiversity hotspot. New Phytol. 215:891-905.

Phillimore, A.B., C.D.L. Orme, G.H. Thomas, T. M. Blackburn, P.M. Bennett, K.J. Gaston, and I.P.F. Owens. 2008. Sympatric speciation in birds is rare: insights from range data and simulations. Am. Nat. 171:646-657.

Pickrell, J.K. and J.K. Pritchard. 2012. Inference of population splits and mixtures from genome-wide allele frequency data. PLoS Genet 8:e1002967.

Potvin, D.A., K.M. Parris, and R.A. Mulder. 2011. Geographically pervasive effects of urban noise on frequency and syllable rate of songs and calls in silvereyes (*Zosterops lateralis*). Proc. R. Soc. B. 278:2464-2469.

Potvin, D.A., and K.M. Parris. 2012. Song convergence in multiple urban populations of silvereyes (*Zosterops lateralis*). Ecol. Evol. 2:1977-1984.

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

Price, T.D. 2008. Speciation in birds. Sinauer Associates, Sunderland, MA.

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 toolset for whole-genome association and population-based linkage analysis. Am. J. of Hum. Genet. 81:559-575.

R Development Core Team. 2016. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/.

Raj, A., M. Stephens, and J.K. Pritchard. 2014. fastSTRUCTURE: Variational inference of population structure in large SNP data sets. Genetics 197:573-189.

Reich, D., K. Thangaraj, N. Patterson, A.L. Price, and L. Singh. 2009. Reconstructing Indian population history. Nature 461:489-494.

Rhymer, J.M., and D. Simberloff. 1996. Extinction by hybridization and introgression. Annu. Rev. Ecol. Syst. 27:83-109.

Robertson, A., A.C. Newton, and R.A. Ennos. 2004. Multiple hybrid origins, genetic diversity and population genetic structure of two endemic *Sorobus taxa* on the Isle of Arran, Scotland. Mol. Ecol. 13:123-134.

Rodríguez-Gómez, F., and J.F. Ornelas. 2013. Genetic, phenotypic, and ecological divergence with gene flow at the Isthmus of Tehuantepec: the case of the azure-crowned hummingbird (*Amazilia cyanocephala*). J. Biogeogr 40:1360-1373.

Rodríguez-Gómez, F., and J.F. Ornelas. 2014. Genetic divergence of the Mesoamerican azure-crowned hummingbird (*Amazilia cyanocephala*, Trochilidae) across the Motagua-Polochic-Jocotán fault system. J. Zool. Syst. Evol. Res. 52:142-153.

Rodríguez-Gómez, F., and J.F. Ornelas. 2015. At the passing gate: past introgression in the process of species formation between *Amazilia violiceps* and *A. viridifrons* hummingbirds along the Mexican transition zone. J. Biogeogr 42:1305-1318.

Rodríguez-Gómez, F., and J.F. Ornelas. 2018. Genetic structuring and secondary contact in the white-chested Amazilia hummingbird species complex. J. Avian Biol 49:jav-01546.

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

Sætre, G-P., and S.A. Sætre. 2010. Ecology and genetics of speciation in the *Ficedula* flycatchers. Mol. Ecol. 19:1091-1106.

Sandvig, E.M., T. Coulson, J. Kikkawa, and S.M. Clegg. 2017. The influence of climactic variation and density on the survival of an insular passerine *Zosterops lateralis*. PLoS ONE 12:e0176360.

Sardell, J.M., and J.A.C. Uy. 2016. Hybridization following recent secondary contact results in asymmetric genotypic and phenotypic introgression between island species of *Myzomela* honeyeaters. Evolution 70:257-269.

Savolainen, V., M-C. Anstett, C. Lexer, I. Hutton, J.J. Clarkson, M.V. Norup, M.P. Powell, D. Springate, N. Salamin, and W.J. Baker. 2006. Sympatric speciation in palms on an oceanic island. Nature 441:210-213.

Schneider, C.A., W.S. Rasband, and K.W. Eliceiri. 2012. NIH Image to ImageJ: 25 years of image analysis. Nat. Methods 9:671-675.

Schreiber K.J., and J. Lancho Rojas. 1995. The Puquios of Nasca. Lat. Am. Antiq. 6:229-154.

Schulenberg, T.S., D.F. Stotz, D.F. Lane, J.P. O'Neill, and T.A. Parker III. 2010. Birds of Peru: revised and updated edition. Princeton University Press, Princeton, NJ.

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.

Searcy, W.A. 1990. Species recognition of song by female red-winged blackbirds. Animal Behav. 40:1119-1127.

Seddon, N., C.A. Botero, J.A. Tobias, P.O. Dunn, H.E.A. MacGregor, D.R. Rubenstein, J.A.C. Uy, J.T. Weir, L.A. Whittingham, and R.J. Safran. 2013. Sexual selection accelerates signal evolution during speciation in birds. Proc. R. Soc. B. 280:20131065.

Servedio, M.R., and M.A.F. Noor. 2003. The role of reinforcement in speciation: theory and data. Annu. Rev. Ecol. Evol. Syst. 34:339-364.

Shendure, J., and H. Ji. 2008. Next-generation DNA sequencing. Nat. Biotechnol. 26:1135-1145.

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, B.T., and J. Klicka. 2013. Examining the role of effective population size on mitochondrial and multilocus divergence time discordance in a songbird. PLoS ONE 8:e55161.

Sorenson, M.D., J.C. Ast, D.E. Dimcheff, T. Yuri, and D.P. Mindell. 1999. Primers for a PCR-based approach to mitochondrial genome sequencing in birds and other vertebrates. Mol. Phylogenet. Evol. 12:105-114.

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

Stamatakis, A., P. Hoover, and J. Rougemont. 2008. A rapid bootstrap algorithm for the RAxML web servers. Syst. Biol. 57:758-771.

Steeves, T.E., R.F. Maloney, M.L. Hale, J.M. Tylianakis, and N.J. Gemmell. 2010. Genetic analyses reveal hybridization but no hybrid swarm in one the world's rarest birds. Mol. Ecol. 19:5090-5100.

Stein, A.C., and J.A.C. Uy. 2006. Unidirectional introgression of a sexually selected trait across an avian hybrid zone: a role for female choice? Evolution 60:1476-1485.

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

Tajima, F. 1989b. The effect of change in population size on DNA polymorphism. Genetics 123:597-601.

Taylor, S.A., R.L. Curry, T.A. White, V. Ferretti, and I. Lovette. 2014. Spatiotemporally consistent genomic signatures of reproductive isolation in a moving hybrid zone. Evolution 69:3066-3081.

Toews, D.P. and A. Brelsford. 2012. The biogeography of mitochondrial and nuclear discordance in animals. Mol. Ecol. 21:3907-3930.

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.

Turner, S.D. 2014. qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. bioRxiv DOI: 10.1101/005165.

Urban, M.C. 2015. Accelerating extinction risk from climate change. Science 348: 571-573

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

Uy, J.A.C., and R.J. Safran. 2013. Variation in the temporal and spatial use of signals and its implications for multimodal communication. Behav. Ecol. Sociobiol. 67:1499-1511.

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

van Balen, B. 2001. Family Zosteropidae (White-eyes). In Handbook of the birds of the world, Vol 13: Penduline-tits to Shrikes. Eds: J. del Hoyo, A. Elliott, D.A. Christie. Lynx Edicions, Barcelona, Spain.

Vorobyev M., and D. Osorio. 1998. Receptor noise as a determinant of colour thresholds. Proc. Biol. Sci. 265:351-358.

Wang, J. 2005. Estimation of effective population sizes from data on genetic markers. Phil. Trans. R. Soc. B 360:1395-1409.

Wang, J., E. Santiago, and A. Caballero. 2016. Prediction and estimation of effective population size. Heredity 117:193-206.

Waples, R.S., and C. Do. 2008. LDNE: a program for estimating effective population size from data on linkage disequilibrium. Mol. Ecol. Resour. 8:753–756.

Warren, B.H., E. Bermingham, R.P. Prys-Jones, and C. Thébaud. 2006. Immigration, species radiation and extinction in a highly diverse songbird lineage: white-eyes on Indian Ocean islands. Mol. Ecol. 15:3769-3786

Webb, W.C., J. M. Marzluff, and K.E. Omland. 2011. Random interbreeding between cryptic lineages of the Common Raven: evidence for speciation in reverse. Mol. Ecol. 20:2390-2402.

Weber, E.H. 1834. De pulsu, resorpitione, auditu et tactic: annotationes anatomicae et physiologicae. Leipzig: Koehlor.

Weir, J.T., and M. Price. 2011. Andean uplift promotes lowland speciation through vicariance and dispersal in *Dendrocincla* woodcreepers. Mol. Ecol. 20:4550-4563.

Weeks, B.C., and S. Claramunt. 2014. Dispersal has inhibited avian diversification in Australasian archipelagoes. Proc. R. Soc. B. 281:20141257.

Weeks, B.C., J. Diamond, P.R. Sweet, C. Smith, G. Scoville, T. Zinghite, and C.E. Filardi. 2017. New behavioral, ecological, and biogeographic data on the montane avifuana of Kolombangara, Solomon Islands. Wilson J. Ornithol. 129:676-700.

Weller, A-A. 2000. Biogeography, geographic variation and habitat preference in the Amazilia Hummingbird, *Amazilia amazilia* Lesson (Aves: Trochilidae), with notes on the status of *Amazilia alticola* Gould. J. Ornithol. 141:93-101.

Weller, A.A., G.M. Kirwan, and P. Boesman. 2019. Amazilia Hummingbird (*Amazilia amazilia*). In: del Hoyo, J., A. Elliott, J. Sargatal, D.A. Christie, and E. de Juana. (eds.). Handbook of the birds of the world alive. Lynx Edicions, Barcelona.

Whaley, O.Q., A. Orellana-Garcia, and J.O. Pecho-Quispe. 2019. An annotated checklist to vascular flora of the Ica Region, Peru—with notes on endemic species, habitat, climate, and agrobiodiversity. Phyotaxa 389:1-125.

While, G.M., S. Michaelides, R.J.P. Heathcote, H.E.A. MacGregor, N. Zajac, J. Beninde, P. Carazo, G. Pérez i de Lanuza, R. Sacchi, M.A.L. Zuffi, et al. 2015. Sexual selection drives asymmetric introgression in wall lizards. Ecol. Lett. 18:1366-1375.

Whitmore, T.C. 1969. The vegetation of the Solomon Islands. Phil. Trans. Royal Soc. B 255: 259-270.

Wickramasinghe, N., V.V. Robin, U. Ramakrishnan, S. Reddy, and S.S. Seneviratne. 2017. Non-sister Sri Lankan white-eyes (genus *Zosterops*) are a result of independent colonizations. PLoS ONE 12:e0181441.

Williams, S.E., E.E. Bolitho, and S. Fox. 2003. Climate change in Australian tropical rainforests: an impending environmental catastrophe. Proc. R. Soc. Lond. B 270:1887-1892.

Wolfe, K.H., and Li, W-H. 2003. Molecular evolution meets the genomics revolution. Nat. Genet. 33:255-265.

Woolfit, M., and L. Bromham. 2005. Population size and molecular evolution on islands. Proc. R. Soc. B. 272:2277-2282.

Wright, S. 1931. Evolution in Mendelian populations. Genetics 16:97–159.

Zarza, E., B.C. Faircloth, W.L.E. Tsai, R.W. Bryson Jr., J. Klicka, and J.E. McCormack. 2016. Hidden histories of gene flow in highland birds revealed with genomic markers. Mol. Ecol. 25:5144-5157.