**Phylogenomics of Avian Taxa in The Southwest Pacific**

By

Lukas B. Klicka

Submitted to the graduate degree program in Ecology and Evolutionary Biology and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Chair: Robert G. Moyle

Mark T. Holder

Rich E. Glor

Andrew E. Z. Short

Robert L. Unckless

Date Defended: 31 January 2020

The dissertation committee for Lukas B. Klicka certifies that this is the approved version of the following dissertation:

# Phylogenomics of Avian Taxa in The Southwest Pacific

Chairperson: Robert G. Moyle

Date Approved: 7 February 2020

## Abstract

Island archipelagos provide natural laboratories to investigate how geographical and ecological differences impact the process of speciation. I investigated the evolutionary history of four different bird species complexes distributed throughout the southwest Pacific. Within the Solomon Islands, all species complexes contained genetic evidence for independent sister lineages on Makira and Ugi. All four systems also identified a close relationship between populations on Guadalcanal, Isabel, and Choiseul, with only one system indicating the potential for multiple lineages across these islands. For three of the species complexes, I uncovered evidence potentially indicating geneflow between distinct lineages, suggesting differentiation has proceeded in the presence of geneflow. This dissertation contributes to a growing body of literature investigating genetic diversity and genomic differentiation for taxa across the southwest Pacific.

# Acknowledgments

I consider myself very fortunate to have been able to pursue my doctoral degree at the University of Kansas (KU). I have been surrounded by mentors and friends for the past six years at KU, and this group of people has contributed greatly to my success and development as both a scientist and an educator. I would first like to thank Dr. Robert Moyle. In addition to being a supportive mentor for my research, Rob unfailingly supported my pursuit of opportunities to further develop myself as an educator. In fact, Rob provided me with the unique opportunity to mentor several undergraduate students as part of a field-based course in the Solomon Islands. Furthermore, Rob is a proponent of developing a healthy work-life balance and in my first year at KU, he introduced me to the world of waterfowl hunting. The time I have spent in a duck blind has produced some of my fondest memories at KU, and I am grateful to Rob for introducing me to the sport.

I would also like to thank my fellow Moyle Lab labmates, both past and present: Dr. Joseph Manthey, Dr. Carl Oliveros, Luke Campillo, Jack Hruska, Jesse Holmes, Lucas Decicco, Emily Ostrow, and Devan DeRaad. Specifically, I want to thank Joe and Carl for their immediate guidance upon starting my program at KU. I consider them both friends and mentors, and my development—particularly with respect to bioinformatics and next-generation sequencing techniques—would have been greatly delayed without their patient assistance. In addition to the mentorship I have received from within the Moyle Lab, I am also thankful for the members of my academic committee: Dr. Rich Glor, Dr. Mark Holder, Dr. Andrew Short, Dr. Rob Unckless, and Dr. Stephen Benedict. I am thankful for the time and effort they dedicated to reviewing my proposals and my dissertation, but also for their availability whenever I asked for assistance. My committee was a supportive group that consistently provided ideas and direction to help me

achieve my personal career goals, and I am thankful for their mentorship. Beyond my committee and lab mates, I have been fortunate to receive continued friendship and guidance from my former master's degree advisor Dr. Kevin Burns long after my departure from San Diego State University. Further, thank you to Dr. Leo Smith. I was able to discuss any topic with Leo, personally or professionally, and I am grateful to have had his advice and input. I would also like to thank Dr. Kirsten Jensen for her role in my professional development. I appreciate her time and guidance. Thank you to Dr. Alana Alexander. During my tenure at KU, Alana was a close friend, and helped me to solve my bioinformatic issues.

I would not be the educator I am today without the influence of many individuals at KU and therefore I would like to thank Dr. Lena Hileman, Dr. Paulyn Cartwright, Dr. Leo Smith, Dr. Rob Unckless (Molecular Biosciences), Dr. Rob Ward (Molecular Biosciences), Dr. Maria Orive, Dr. Kris Holder, Dr. Jamie Walters, Dr. Rob Moyle, Dr. Mark Mort, and Dr. Justin Blumenstiel. I have learned a great deal about teaching from, and have appreciated my time working with, each of them. Furthermore, I would like to thank the students and staff of Haskell Indian Nations University. Teaching at Haskell was an incredible opportunity for me, and I am grateful to those that made it such an enjoyable experience.

Thanks are also due to the support staff of Ecology and Evolutionary Biology department and University of Kansas Natural History Museum. In particular, thank you to Aagje Ashe, Jaime Keeler, and Lori Schlenker. There was never a question they could not answer nor paperwork with which they could not assist and I am thankful to have had their help during my tenure at KU.

I want to broadly thank the graduate students of the Ecology and Evolutionary Biology department, and especially thank Dr. Joe Manthey, Luke Campillo, Lucas Decicco, Jack Hruska,

Dr. Carl Oliveros, Stephen Baca, Matt Girard, Dr. Patrick Monnahan, Dr. Kaila Colyott, Kevin Chovanec, Dr. Scott Travers, Karen Olson, Mark Herr, Robin Abraham, Ali Khalighifar, Fernando Machado-Stredel, and Jacob Hopkins for their friendship and academic discussions. Furthermore, thank you to Kaylee Herzog. I could not have asked for a more supportive, thoughtful, and loving partner. I am truly fortunate to have had her by my side these past four years.

The research I conducted at KU depended on intensive investment in fieldwork, and therefore would not have been possible without broad collaborations between researchers, museums, local communities, and governments. Thank you to Dr. Chris Filardi, Dr. Catherine Smith, Dr. Leo Joseph, Dr. John Dumbacher, Dr. Albert Uy, and Dr. Rob Moyle for their longstanding investment of time and resources to completing field work across the Indo-Pacific. Thank you to the museums and support staff of the Australian National Wildlife Collection, University of Washington Burke Museum, University of Miami, California Academy of Sciences, University of Kansas Natural History Museum, American Museum of Natural History and Western Australian Museum for supporting the field studies and providing the tissue samples necessary for my research.

My personal field research was conducted in the Solomon Islands, and so I would like to thank the Solomon Islands' Ministry of Environment, Climate Change, Disaster Management, and Meteorology. A special thank you also to the people and communities of the Solomon Islands for welcoming me and working with me in the field. I was fortunate to spend my field research with some amazing colleagues and therefore thank you Luke Campillo, Dr. Scott Travers, David Bosetto, Ikuo Gumo Keanu Tígulu, and Karen Olson.

This research was made possible through funding from several sources including the National Science Foundation (DEB 1557053 and DEB 1557051), University of Kansas Biodiversity Institute, University of Kansas Department of Ecology and Evolutionary Biology, Center for Molecular Analysis of Disease Pathways–Centers of Biomedical Research excellence (CMADP-COBRE), and the Society of Systematic Biologists (SSB).

Lastly, I want to thank my friends and family that are spread across the country. I always look forward to our annual family events like the Fourth of July and Christmas, and I am very grateful to those friends that have been able to make a trip to Kansas or have been able to stay in touch over the years. It has meant a lot. I am truly grateful I decided to continue my education at KU. It has been an exciting and formative chapter in my life.

Table of Contents

## List of Figures



## List of Tables

## Introduction

In animal systems, the prevailing paradigm for speciation to occur is geographic isolation. Islands and island systems present an ideal geographic landscape to study animal speciation because of the naturally disjunct distributions they enforce on terrestrial organisms. Beyond fragmented landscapes, many factors that are expected to influence speciation vary between islands such as age, degree of isolation, size, elevation, and climate. Thus, island systems have often drawn the interest of evolutionary biologists and have had a long history of shaping evolutionary theory. In particular, the southwest Pacific contains an abundance of island archipelagos in which to study the process of diversification and has been a focal region for speciation research for decades. Within the diverse and fragmented landscape of the southwest Pacific, one archipelago, the Solomon Islands, encapsulates many of the advantages for investigating diversification in island taxa. The archipelago consists of 6 major islands, some of which were connected via Pleistocene land bridges (e.g. Bougainville, Choiseul, and Isabel), whereas others (Malaita and Makira) have remained geographically isolated over geologic history. On an even smaller geographic scale, the New Georgia Group islands in Western Province, represent an archipelago within the Solomon Islands. The New Georgia Group as a whole has remained disconnected from the major landmasses within the Solomon Islands, but within the group, periods of low sea levels have connected some, but not all, of the islands. Therefore, lineages that have dispersed across the Solomon Islands can be examined over a broad spectrum of geographic isolation.

The goal of this dissertation was to reconstruct the evolutionary history of avian taxa distributed throughout the Solomon Islands. In doing so, I endeavored to investigate the processes influencing the patterns of genetic differentiation. For this research, I identified genetic

loci from throughout the genome using two reduced-representation genomic approaches and their associated bioinformatic pipelines. For the Rufous Fantail (*Rhipidura rufifrons*, Chapter 1) and the White-bellied Cuckooshrike (*Coracina papuensis*, Chapter 2) I used restriction site associated DNA sequencing and targeted sequence capture of ultra-conserved elements for Monarch flycatchers (Monarchidae, Chapter 3).

For the Rufous Fantail in Chapter 1, I used thousands of single nucleotide polymorphisms from throughout the genome to identify distinct evolutionary lineages and examine patterns of genetic diversity. I found morphological divergence to be a poor approximation of genetic divergence, did not find genetic evidence to support some named Solomon Island taxa as independent lineages, and uncovered evidence for gene flow across hundreds of kilometers of open ocean.

Within the Solomon Islands and in the southwest Pacific in general, avian phylogeographers have mostly focused on species with low dispersal ability. In Chapter 2, I produced one of the first genomic datasets for a putatively high dispersal species and investigated the genetic diversity and patterns of differentiation for the White-bellied Cuckooshrike. Phylogenetic analyses of genome-wide variation revealed several well-supported, independent lineages of island taxa, but limited support was identified for multiple lineages within Australia. Additionally, I identified multiple individuals with substantially admixed genetic backgrounds. Thus, it appears lineage differentiation for some White-bellied Cuckooshrike populations has continued in the presence of gene flow.

For Chapter 3, I investigated phylogeographic patterns in two species complexes within the Monarchidae (Chestnut-bellied and Pied Monarchs). In order to include specimens for which only historical DNA was available I used a targeted sequence capture method to sequence a few

thousand loci from throughout the genome. Despite similar life histories, distributions and lineage age, the Pied Monarchs have differentiated across the Solomon Islands into more described taxa than the Chestnut-bellied Monarchs. For the Pied Monarchs, the samples for every large island (i.e., Bougainville, Choiseul, Isabel, Guadalcanal, Malaita, and Makira) showed some structure, with individuals from each island forming a clade and multiple lineages identified within the New Georgia Group.

This dissertation contributes to a growing body of literature investigating genomic evidence for genetic diversity and differentiation for taxa across the southwest Pacific. This research will provide a foundation for future studies investigating the maintenance and formation of population differentiation in the presence of gene flow.

# Chapter 1: Genomic and geographic diversification of a "great-speciator" (*Rhipidura rufifrons*)

## Introduction

Island systems have a long history of influencing the development of evolutionary theory (Darwin, 1859; Mayr, 1942; Wallace, 1881). Island archipelagos vary in many aspects that are expected to influence biological diversification such as age, degree of isolation, size, elevation, and climate. This variation within and among archipelagos provides natural laboratories in which to test hypotheses of how geographical and ecological differences influence the process of diversification (MacArthur & Wilson, 1967). Given the abundant number of archipelagos and thousands of islands within the southwest Pacific, it is unsurprising that early speciation research focused extensively on terrestrial biodiversity in this region (Diamond, 1974; Diamond, Gilpin, & Mayr, 1976; MacArthur & Wilson, 1963, 1967; Mayr, 1942; Mayr & Diamond, 2001). Known for morphologically diverse lineages, "great speciators" (Mayr & Diamond, 2001) represent a particularly interesting evolutionary phenomenon in which avian species complexes appear to have diversified rapidly across the southwest Pacific (Andersen, Oliveros, Filardi, & Moyle, 2013; Andersen et al., 2015; Irestedt et al., 2013; Moyle, Filardi, Smith, & Diamond, 2009; Pedersen, Irestedt, Joseph, Rahbek, & Jønsson, 2018). Rapid radiations are not limited to Pacific island systems (Campagna, Gronau, Silveira, Siepel, & Lovette, 2015; Koblmüller, Egger, Sturmbauer, & Sefc, 2010; Losos, Jackman, Larson, de Queiroz, & Rodríguez-Schettino, 1998; Rees, Emerson, Oromí, & Hewitt, 2001), but because many rapid radiations in the southwest Pacific share broadly overlapping distributions, they naturally lend themselves to detailed investigations of biogeographic theory in a comparative framework.

Rapid geographic radiations often comprise species complexes with uncertain taxonomic divisions or phylogenetic relationships. Recent phylogenetic work, however, has begun to unravel some of the mystery previously surrounding the evolution of these taxa. For example,

*Todiramphus chloris*, a species complex distributed over 16,000 km and containing 50 subspecies, was recently examined using modern phylogenetic methods (Andersen et al., 2015). Strikingly, despite already comprising 50 described taxa, 10 additional *Todiramphus* species were found to be embedded within the *T. chloris* complex. Although taxonomically interesting in and of itself, the comprehensive phylogeny of the *T. chloris* complex also uncovered biogeographic patterns warranting further exploration. For example, two distinct lineages of the Mariana Islands did not form sister clades and displayed divergent biogeographic patterns. One lineage from the Mariana Islands grouped with Polynesian and Micronesian lineages, and the second was recovered as sister to only Melanesian populations. A similar pattern of divergent lineage relationships within an archipelago was identified within populations of *T. chloris* from the Solomon Islands, where the populations of Makira and Ugi in the southeast Solomon Islands were more closely related to Polynesian populations in Vanuatu, Fiji, American Samoa and Tonga than to individuals from elsewhere in the Solomon Islands.

Conversely, among populations of *Pachycephala pectoralis* representing a separate rapid radiation, a Solomon Islands clade was found to be sister to Australian, Melanesian, and Polynesian radiations rather than embedded within these regions, as was found in *T. chloris* (Andersen et al., 2014; Jønsson et al., 2014). Furthermore, despite the geographic proximity of the Louisiade Archipelago to Papua New Guinea, populations of *P. pectoralis* from the Louisiade Archipelago were not closely related to populations from mainland New Guinea, but rather sister to the entire species complex, which spans the Australasian region (Andersen et al., 2014). These and other explorations of rapid radiations in the Indo-Pacific (e.g., Andersen et al., 2013; Moyle et al., 2009; Pedersen et al., 2018) have been based mostly on relatively small genetic datasets (< 10 loci). In this study we leverage newer sequencing technologies to enable

exploration of a radiation using a genome-wide dataset. Such data-rich approaches can help to resolve complex biogeographic relationships and population demographic histories that smaller genetic datasets are less likely to inform.

The *Rhipidura rufifrons* species complex is a widespread, phenotypically diverse (Pratt, 2010) geographic radiation that is broadly sympatric with the aforementioned species complexes. *Rhipidura rufifrons* is currently composed of 18 described subspecies (Clements et al., 2019) distributed predominately throughout Melanesia, with additional taxa from Micronesia and Australia. *Rhipidura rufifrons* inhabits all of the major islands of the Solomon Islands, an archipelago that encapsulates the advantages of studying island systems, but also introduces the possibility of Pleistocene land bridge connections between some islands. In addition, populations of *R. rufifrons* reside on remote island archipelagos that are isolated by hundreds or even thousands of kilometers of open ocean from the nearest *R. rufifrons* population. Therefore, *R. rufifrons* populations enable exploration of diversification over a broad spectrum of geographic isolation.

Prior molecular phylogenetic work on this complex has either informed the phylogenetic position of *R. rufifrons* within the family Rhipiduridae (Nyári, Benz, Jønsson, Fjeldså, & Moyle, 2009), or concentrated on a limited geographic subset of *R. rufifrons* populations (Weidemann, 2015). A distribution-wide genomic analysis has yet to be completed for this complex. *Rhipidura rufifrons* subspecies display varied levels of morphological divergence (Pratt, 2010). For example, *R. r. saipanensis* exhibits little plumage differentiation compared to the nominate subspecies, despite being the most geographically isolated population. Conversely, *R. r. ugiensis* and *R. r. russata* differ distinctly from one another in plumage coloration but are separated by less than ten kilometers on Ugi Island and Makira Island, respectively. Given the wide

distribution of subspecies in this complex, their rapid diversification, and their morphological variation, *R. rufifrons* is an ideal system in which to both broaden our understanding of evolution during rapid radiations and address biogeographic hypotheses proposed by prior investigations. Therefore, the objectives for this study are threefold: 1) reconstruct the evolutionary history of the *R. rufifrons* complex, 2) identify biogeographic patterns within the complex, and 3) assess the levels of gene flow in a rapid radiation.

## Methods

### Sampling

Sampling included 94 individuals representing 19 named taxa from five species: *Rhipidura rufifrons* (12 taxa)*, R. dryas* (4 taxa), *R. lepida*, *R. teysmanni*, and *R. dahli* with the last three species included as outgroups based on Nyári et al. (2009) and unpublished mitochondrial DNA data (Table 1). Three described subspecies within *R. rufifrons* (*R. r. torrida* - Moluccas, *R. r. versicolor* - West Caroline Island, and *R. r. uraniae* - Guam) currently lack modern sampling and unfortunately *R. r. uraniae* is extinct (Boles & Christi, 2019). Furthermore, Mayr and Diamond (2001) included *R. semirubra* as a subspecies within the *R. rufifrons* complex, whereas other authorities split the Manus Island endemic as its own species (Clements et al., 2019). Nevertheless, modern samples are not available for *R. semirubra*, and thus, this taxon was not included in this study. We extracted genomic DNA from blood or tissue samples using a QIAGEN DNeasy blood and tissue kit for all individuals, and quantified DNA concentrations with a Qubit Fluorometer 2.0 (Life Technologies).

### Sequencing and Bioinformatics

We performed a single digest RAD-seq protocol (Miller, Dunham, Amores, Cresko, & Johnson, 2007) to obtain thousands of loci from across the *R. rufifrons* genome. We followed

procedures outlined by Manthey, Campillo, Burns, and Moyle (2016) to generate the DNA libraries. We used NdeI to digest the genomic DNA and ligated custom barcoded adapters (Andolfatto et al., 2011) to permit the multiplexing of many individuals. We size selected fragments in a range between 450–600 bp using a Pippin Prep (Sage Science) electrophoresis cassette. Samples were sequenced using partial lanes of three different sequencing runs on an Illumina HiSeq2500 and an Illumina NextSeq 550 for 100 bp single-end reads at the University of Kansas Genome Sequencing Core Facility.

We used the STACKS v2.3 (Catchen, Amores, Hohenlohe, Cresko, & Postlethwait, 2011) pipeline to assemble loci and produce a single nucleotide polymorphism (SNP) dataset from the sequencing data. Individuals were de-multiplexed and low-quality reads were removed using the *process_RADtags* script from STACKS. Following de-multiplexing, we ran the modules *ustacks*, *cstacks*, and *sstacks*. We used the *ustacks* module to identify loci within an individual initially using default parameters for number of mismatches allowed between stacks (-M 2) and the number of reads required to build a stack (-m 3). Next, we ran *cstacks* to combine individual loci into a catalogue of loci, permitting three mismatches (-n 3) across individuals. Then, we matched each individual's data to the catalogue with *sstacks* using default parameters. We transposed the dataset using *tsv2bam* and aligned and called SNPs using the *gstacks* module. Using the *populations* module within STACKS, we filtered out loci with a minor allele frequency below 0.05 and observed heterozygosity above 0.5. In order to examine possible influences of parameter choice on downstream analyses, we re-ran this protocol, iteratively modifying the parameters M (1–4), m (3, 5, 7), and n (1, 3, 5). Because sequence data for this project were amalgamated from three separate Illumina runs, we accounted for library specific

loci by dropping loci present in fewer than 70% of individuals because no single library contributed more than 55% of the total individuals.

Table 1. Locality and voucher information for the individual samples included in chapter 1

| Species | Tissue # | Museum | Locality | Fig 1 Clade | Reads post QC | Cov. Median |
|---|---|---|---|---|---|---|
| *R. r. intermedia* | B31321 | ANWC | Australia: Queensland | I | 3960589 | 9 |
| *R. r. intermedia* | B31559 | ANWC | Australia: Queensland | I | 3217638 | 19 |
| *R. r. intermedia* | B43019 | ANWC | Australia: Queensland | I | 1086917 | 36 |
| *R. r. intermedia* | B57118 | ANWC | Australia: Queensland | I | 2131317 | 22 |
| *R. r. intermedia* | B57131 | ANWC | Australia: Queensland | I | 586575 | 20 |
| *R. r. intermedia* | B57199 | ANWC | Australia: Queensland | I | 1661127 | 22 |
| *R. r. intermedia* | B56097 | ANWC | Papua New Guinea: Western Province | I | 1286979 | 15 |
| *R. r. intermedia* | B56405 | ANWC | Papua New Guinea: Western Province | I | 790126 | 25 |
| *R. r. rufifrons* | B44315 | ANWC | Australia: Queensland | I | 922775 | 29 |
| *R. r. rufifrons* | B46863 | ANWC | Australia: New South Wales | I | 2891189 | 24 |
| *R. r. rufifrons* | B49720 | ANWC | Australia: New South Wales | I | 2629858 | 17 |
| *R. r. louisiadensis* | 97870 | CAS | Papua New Guinea: Duchess Is | II | 793509 | 22 |
| *R. r. louisiadensis* | 97887 | CAS | Papua New Guinea: Panamoti Is | II | 1344191 | 20 |
| *R. r. louisiadensis* | 97970 | CAS | Papua New Guinea: Panamoti Is | II | 3493235 | 26 |

| Species | Tissue # | Museum | Locality | Fig 1 Clade | Reads post QC | Cov. Median |
|---|---|---|---|---|---|---|
| *R. r. louisiadensis* | 184548 | BPBM | Papua New Guinea: Duchess Is | II | 1740924 | 27 |
| *R. r. louisiadensis* | 96782 | CAS | Papua New Guinea: Panapompom Is | II | 1290250 | 14 |
| *R. r. louisiadensis* | 96783 | CAS | Papua New Guinea: Haszard Is | II | 1758740 | 11 |
| *R. r. louisiadensis* | 96785 | CAS | Papua New Guinea: Hummock Is | II | 1668508 | 11 |
| *R. r. agilis* | 19407 | KUNHM | Solomon Is: Santa Cruz | III | 895692 | 15 |
| *R. r. agilis* | 19408 | KUNHM | Solomon Is: Santa Cruz | III | 3959227 | 19 |
| *R. r. agilis* | 19417 | KUNHM | Solomon Is: Santa Cruz | III | 1234707 | 14 |
| *R. r. saipanensis* | 22578 | KUNHM | Northern Mariana Is: Saipan | III | 1513336 | 12 |
| *R. r. saipanensis* | 22588 | KUNHM | Northern Mariana Is: Saipan | III | 1611507 | 13 |
| *R. r. saipanensis* | 22597 | KUNHM | Northern Mariana Is: Tinian | III | 3561756 | 20 |
| *R. r. saipanensis* | 22601 | KUNHM | Northern Mariana Is: Tinian | III | 689131 | 11 |
| *R. r. ugiensis* | 15928 | KUNHM | Solomon Is: Ugi | V | 926692 | 12 |
| *R. r. ugiensis* | M007 | U of Miami | Solomon Is: Ugi | V | 3216729 | 17 |
| *R. r. ugiensis* | M008 | U of Miami | Solomon Is: Ugi | V | 2805477 | 18 |
| *R. r. kuperi* | M105 | U of Miami | Solomon Is: Santa Catalina | VI | 1950950 | 15 |
| *R. r. kuperi* | M106 | U of Miami | Solomon Is: Santa Catalina | VI | 5448399 | 21 |
| *R. r. kuperi* | M109 | U of Miami | Solomon Is: Santa Catalina | VI | 2630954 | 17 |
| *R. r. russata* | M112 | U of Miami | Solmon Is: Makira | VI | 1693946 | 16 |
| *R. r. russata* | M113 | U of Miami | Solmon Is: Makira | VI | 3815039 | 17 |
| *R. r. russata* | 12828 | KUNHM | Solmon Is: Makira | VI | 661948 | 15 |

| Species | Tissue # | Museum | Locality | Fig 1 Clade | Reads post QC | Cov. Median |
|---|---|---|---|---|---|---|
| *R. r. russata* | 12832 | KUNHM | Solmon Is: Makira | VI | 774385 | 13 |
| *R. r. russata* | 13547 | KUNHM | Solmon Is: Makira | VI | 639724 | 9 |
| *R. r. russata* | 15915 | KUNHM | Solmon Is: Makira | VI | 675950 | 10 |
| *R. r. granti* | 32098 | KUNHM | Solomon Is: Gizo | VII | 740113 | 12 |
| *R. r. granti* | 32104 | KUNHM | Solomon Is: Gizo | VII | 935143 | 8 |
| *R. r. granti* | 33777 | KUNHM | Solomon Is: Ranongga Is | VII | 1317088 | 18 |
| *R. r. granti* | 33786 | KUNHM | Solomon Is: Ranongga Is | VII | 6060119 | 12 |
| *R. r. granti* | 33791 | KUNHM | Solomon Is: Ranongga Is | VII | 176950 | 11 |
| *R. r. granti* | 33862 | KUNHM | Solomon Is: Vella Lavella | VII | 13321108 | 13 |
| *R. r. granti* | 33910 | KUNHM | Solomon Is: Kolombangra | VII | 2849754 | 10 |
| *R. r. granti* | 33919 | KUNHM | Solomon Is: Kolombangra | VII | 12780508 | 10 |
| *R. r. granti* | Bu66046 | UWBM | Solomon Is: New Georgia | VII | 1249487 | 10 |
| *R. r. granti* | Bu66083 | UWBM | Solomon Is: New Georgia | VII | 458788 | 9 |
| *R. r. brunnea* | 32766 | KUNHM | Solomon Is: Malaita | VIII | 667794 | 9 |
| *R. r. brunnea* | 32767 | KUNHM | Solomon Is: Malaita | VIII | 923712 | 13 |
| *R. r. brunnea* | 32779 | KUNHM | Solomon Is: Malaita | VIII | 1413435 | 19 |
| *R. r. brunnea* | 32783 | KUNHM | Solomon Is: Malaita | VIII | 1038579 | 12 |
| *R. r. brunnea* | 32791 | KUNHM | Solomon Is: Malaita | VIII | 2014330 | 14 |
| *R. r. commoda* | 5282 | KUNHM | Papua New Guinea: Bougainville | VIII | 507790 | 7 |
| *R. r. commoda* | Bu63074 | UWBM | Solomon Is: Choiseul | VIII | 2624511 | 14 |
| *R. r. commoda* | Bu63198 | UWBM | Solomon Is: Choiseul | VIII | 187338 | 8 |
| *R. r. commoda* | Bu58811 | UWBM | Solomon Is: Isabel | VIII | 1346045 | 12 |

| Species | Tissue # | Museum | Locality | Fig 1 Clade | Reads post QC | Cov. Median |
|---|---|---|---|---|---|---|
| *R. r. commoda* | Bu60315 | UWBM | Solomon Is: Isabel | VIII | 138410 | 8 |
| *R. r. commoda* | 32061 | KUNHM | Solomon Is: Shortland Is | VIII | 619934 | 13 |
| *R. r. commoda* | 32066 | KUNHM | Solomon Is: Shortland Is | VIII | 1428900 | 11 |
| *R. r. rufofronta* | 15904 | KUNHM | Solomon Is: Guadalcanal | VIII | 575676 | 12 |
| *R. r. rufofronta* | 15910 | KUNHM | Solomon Is: Guadalcanal | VIII | 2771728 | 15 |
| *R. r. rufofronta* | 32807 | KUNHM | Solomon Is: Guadalcanal | VIII | 1311981 | 9 |
| *R. r. rufofronta* | 32809 | KUNHM | Solomon Is: Guadalcanal | VIII | 1271126 | 15 |
| *R. r. rufofronta* | 32810 | KUNHM | Solomon Is: Guadalcanal | VIII | 1382894 | 11 |
| *R. r. rufofronta* | 32823 | KUNHM | Solomon Is: Guadalcanal | VIII | 840711 | 27 |
| *R. r. rufofronta* | 32854 | KUNHM | Solomon Is: Guadalcanal | VIII | 962325 | 18 |
| *R. r. rufofronta* | Bu60248 | UWBM | Solomon Is: Guadalcanal | VIII | 835126 | 13 |
| *R. d. dryas** | B57249 | ANWC | Australia: Queensland | NA | 2854774 | 17 |
| *R. d. dryas* | 22743 | KUNHM | Australia: Northern Territory | NA | 515364 | 12 |
| *R. d. dryas* | B29495 | ANWC | Australia: Queensland | NA | 2702340 | 11 |
| *R. d. dryas* | B29660 | ANWC | Australia: Queensland | NA | 2613918 | 12 |
| *R. d. dryas* | B29869 | ANWC | Australia: Queensland | NA | 226833 | 10 |
| *R. d. dryas* | B32683 | ANWC | Australia: Queensland | NA | 2568336 | 11 |
| *R. d. dryas* | B33738 | ANWC | Australia: Northern Territory | NA | 3125113 | 7 |
| *R. d. dryas* | B48642 | ANWC | Australia: Northern Territory | NA | 2312421 | 9 |

| Species | Tissue # | Museum | Locality | Fig 1 Clade | Reads post QC | Cov. Median |
|---|---|---|---|---|---|---|
| *R. d. dryas* | B55074 | ANWC | Australia: Western | NA | 981126 | 11 |
| *R. d. semicollaris* | 22704 | WAM | Indonesia: Lesser Sundas | NA | 1944430 | 17 |
| *R. d. semicollaris* | 23563 | WAM | Indonesia: Lesser Sundas | NA | 1864656 | 16 |
| *R. d. semicollaris* | 23884 | WAM | Indonesia: Lesser Sundas | NA | 4015396 | 15 |
| *R. d. semicollaris* | 24442 | WAM | Indonesia: Lesser Sundas | NA | 1026765 | 11 |
| *R. d. semicollaris* | 24503 | WAM | Indonesia: Lesser Sundas | NA | 5576121 | 11 |
| *R. d. squamata* | 24890 | WAM | Indonesia: Banda Islands | NA | 1415653 | 11 |
| *R. d. sumbensis* | 22842 | WAM | Indonesia: Lesser Sundas | NA | 2227098 | 17 |
| *R. teysmanni* | DOT12566 | AMNH | Indonesia: Sulawesi | NA | 882896 | 13 |
| *R. lepida* | 23628 | KUNHM | Palau | NA | 797225 | 13 |
| *R. lepida* | 23660 | KUNHM | Palau | NA | 798799 | 15 |
| *R. lepida* | 23623 | KUNHM | Palau | NA | 3875182 | 18 |
| *R. dahli* | 5305 | KUNHM | | NA | 3446318 | 15 |
| *R. dahli* | 5313 | KUNHM | | NA | 437175 | 9 |

Specimen clade membership for the maximum likelihood phylogeny are listed in column labeled Fig 1 Clade. Cov Median refers to the median depth of coverage per locus. AMNH American Museum of Natural History, ANWC Australian National Wildlife Collection, KUNHM University of Kansas Natural History Museum, UWBM University of Washington Burke Museum, WAM Western Australian Museum

**Analyses**

We performed phylogenetic analyses on the concatenated dataset of SNPs using maximum likelihood. Prior to concatenation, the alleles from each individual's SNPs were collapsed into a single consensus allele per locus, specifying ambiguity codes in the event of polymorphic sites. Loci were then concatenated for each individual and we used RAxML v8.0.19 (Stamatakis, 2014) to identify phylogenetic relationships among individuals under a GTR+G model of nucleotide substitution. Because the analysis omitted constant sites, we performed an ascertainment bias correction and assessed support using 1000 rapid bootstrap replicates.

Population genetic structure was assessed with DAPC (Discriminate Analysis of Principle Components; (Jombart, Devillard, & Balloux, 2010), within the R (R Core Development Team, 2012) package *adegenet* (Jombart, 2008; Jombart & Ahmed, 2011) and STRUCTURE v2.3.4 (Pritchard, Stephens, & Donnelly, 2000). STRUCTURE uses a predetermined number of populations (K) into which individuals are sorted. We used K values from 1–15 and completed ten independent runs for each value of K. We ran STRUCTURE analyses for 550,000 generations per run, with the first 50,000 MCMC generations taken as burn-in. We used likelihood scores and a ΔK (Evanno, Regnaut, & Goudet, 2005) calculation to determine the most likely number of populations. For DAPC, the most likely number of populations was determined based on the Bayesian Information Criterion. For both DAPC and STRUCTURE, we limited analyses to a single SNP per locus.

We assigned individuals to populations based on STRUCTURE and DAPC and inferred a species tree using TreeMix v1.13 (Pritchard et al., 2000), which allows for genetic exchange between populations that is not explained by the species tree alone. Specifically, we iteratively added migration events until these events explained 0.2% or less of the genetic variation. We

determined nodal support for the species tree by using 500 bootstrap replicates and accounted for possible linkage disequilibrium by completing independent runs using a bootstrapping block size (-k) of 100, 500 and 1000 SNPs.

## Results

### SNP Data

After removing low quality reads, we retained a total of 180,955,081 reads from 94 samples, for a mean of 2.01 million reads per individual (range = 35743–13321108; sd: 2.10 million). The lowest coverage individual, UWBM 85583 (Mariana Islands - Rota), had only 35,743 reads and therefore this sample and four other low-coverage samples were not included in subsequent analyses, bringing the total number of individuals down to 89. We determined that no significant differences existed in population differentiation or population relationships between the different data matrices produced by altering the parameters (-M, -m, -n) within the STACKS pipeline. Therefore, we present only the results from the 70% complete matrix using the parameters -M 2, -m 3, -n 3 containing 5625 loci. For analyses that assume marker independence (i.e., STRUCTURE, TreeMix) a single SNP was retained per locus.

### Population Genetics

Comparison of independent STRUCTURE runs using K = 1 to K = 15 (Figure 1), yielded the highest likelihood score at K = 5. A calculation of ΔK identified K = 2 as the best population model; however, a second peak for ΔK was found for the five-population model (K = 5). The five-population model for STRUCTURE recovered the following populations: Australia (1), Louisiade Archipelago (2), Northern Mariana Islands (3), Greater Bukida (Bougainville, Shortland Islands, Choiseul, and Isabel) with Guadalcanal, Malaita, and the New Georgia Group

(4), and southern Solomon Islands (Santa Catalina, Makira, Ugi) (5). The individuals from Santa Cruz shared a genomic background with those from the Northern Mariana Islands and southern

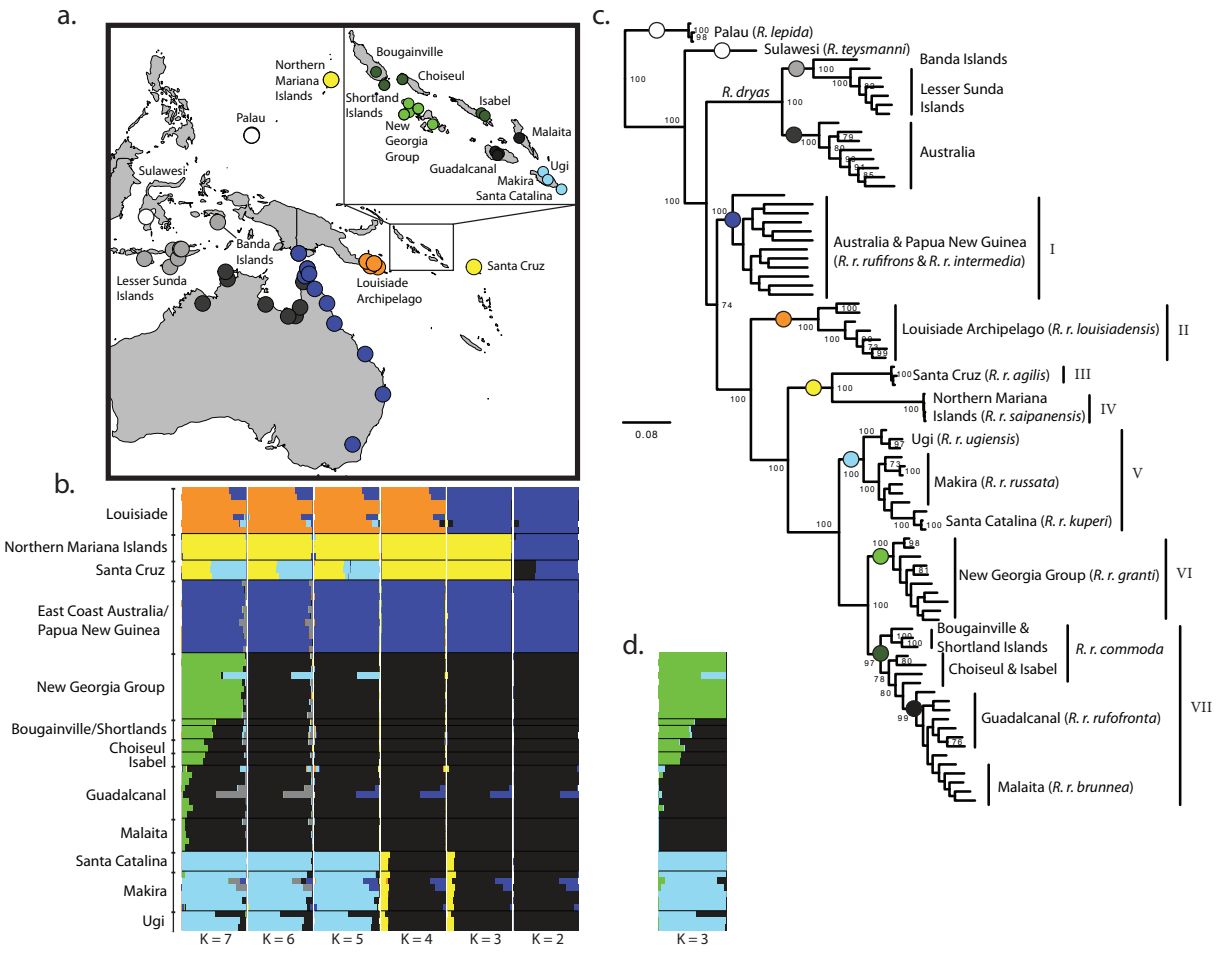


Figure 1. Sampling and genetic structure of the *Rhipidura rufifrons* species complex. a) Sampling locations. Samples of *R. rufifrons* are colored according to their populations assignment for the K = 7 STRUCTURE analysis. b) STRUCTURE results for K = 2–7 using the 70% minimum representation dataset with the outgroup samples removed. c) Maximum likelihood phylogeny obtained using RAxML for the 70% minimum representation SNP dataset with node support determined by rapid bootstrapping and only shown for relationships receiving BS >70. Lineages currently described as *R. rufifrons* are labeled as Clades I–VII. d) The STRUCTURE population model of K = 3 for only Solomon Island individuals.

Solomon Island populations (Figure 1b). Here after, Solomon Islands will be used to refer to the geographic region of the Solomon Archipelago (i.e., Bougainville, Choisel, Isabel, Malaita, Guadalcanal, Makira, etc.). Notably, while the Santa Cruz Islands are included in the Solomon Islands politically, we will discuss them independent of the Solomon Archipelago.

STRUCTURE analyses using only samples from the Solomon Islands supported three populations (K = 3) as the preferred model based on ΔK values and raw likelihood scores (Figure 1d). The three-population model placed the individuals from Makira, Ugi, and Santa Catalina into a cluster. Individuals from Malaita and Guadalcanal individuals formed a second genetic cluster, and individuals from the New Georgia Group formed a third. The Greater Bukida individuals (i.e., those from Bougainville, Shortland Islands, Choiseul, and Isabel) showed varying levels of admixture between the second and third clusters (Figure 1d). A similar result was recovered in the seven-population model (K = 7) using the full dataset (Figure 1b).

DAPC analyses discriminated more population clusters (seven) than STRUCTURE (five) when analyzing all of the *R. rufifrons* taxa. The additional clusters identified by DAPC split Santa Cruz and the North Mariana Islands into distinct clusters, and individuals from the New Georgia Group were grouped as a single population to the exclusion of all other Solomon Island populations (Supporting Information Appendix S1).

Nucleotide diversity within the Solomon Islands ranged from 0.037–0.063 with the majority of the genetic diversity represented by shared polymorphisms (Figure 2). Despite the small size of Ugi and the Santa Catalina Islands, each population from these islands contained similar genetic diversity estimates to the other Solomon Island populations (Figure 2). In contrast, the nucleotide diversity of small but isolated island populations (e.g., Northern Mariana Islands and Santa Cruz) was comparatively low, even with equivalent sample sizes (Figure 2).

**Maximum Likelihood Phylogeny**

Using 11,340 concatenated SNPs from 5625 loci for 89 individuals, we recovered a phylogeny with generally high nodal support. We rooted the phylogeny using *R. dahli* (Nyári et. al. 2009), and we identified seven well-supported lineages within the *R. rufifrons* complex (Figure 1c). *Rhipidura dryas* contained two independent lineages that together formed the sister group to the *R. rufifrons* complex. The oldest relationship in the *R. rufifrons* complex was hypothesized to be between the Australian *R. rufifrons* populations (Clade I; *R. r. rufifrons* and *R. r. intermedia*) and all other *R. rufifrons* lineages. Clade II was composed of individuals from across the Louisiade Archipelago (*R. r. louisadensis*) and it shares a most recent common ancestor with the remaining five well-supported lineages (Clades III–VIII). Two isolated island taxa, *R. r. agilis* (Clade III; Santa Cruz) and *R. r. saipanensis* (Clade IV; Northern Mariana Islands), are sister taxa and together share a most recent common ancestor with all taxa from the Solomon Islands (Clades V–VII). The individuals of the southern Solomon Islands (Clades V; Ugi, Makira, Santa Catalina) shared a sister relationship with individuals from the northern Solomon Islands (Clades VI and VII) and within the northern Solomon Islands, New Georgia Group individuals (Clade VI) were sister to individuals from Bougainville, Shortland Islands, Choiseul, Isabel, Guadalcanal, and Malaita (Clade VII).

**Introgression**

The species tree produced by TreeMix was topologically consistent with the RAxML phylogeny when zero migration edges were permitted. However, with the addition of migration to the

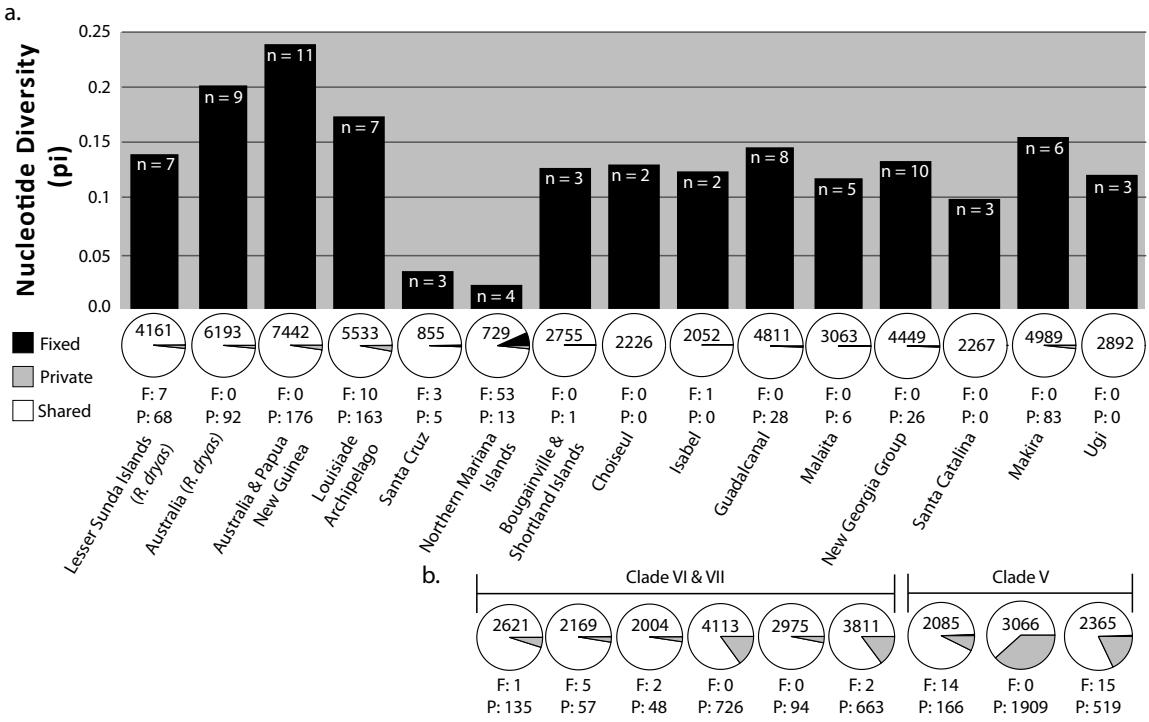


Figure 2. Genetic diversity for *Rhipidura rufifrons* and *R. dryas* populations. a) Nucleotide diversity for each population is shown by the bar graph with sample sizes (n). The proportion of private alleles (P), fixed differences (F), and shared polymorphisms are shown below the bar graph with pie charts. Numbers within or below pie charts indicate counts for private alleles, fixed differences, and shared polymorphisms. b) Private, fixed and shared polymorphisms for Clade V and for Clades VI + VII.

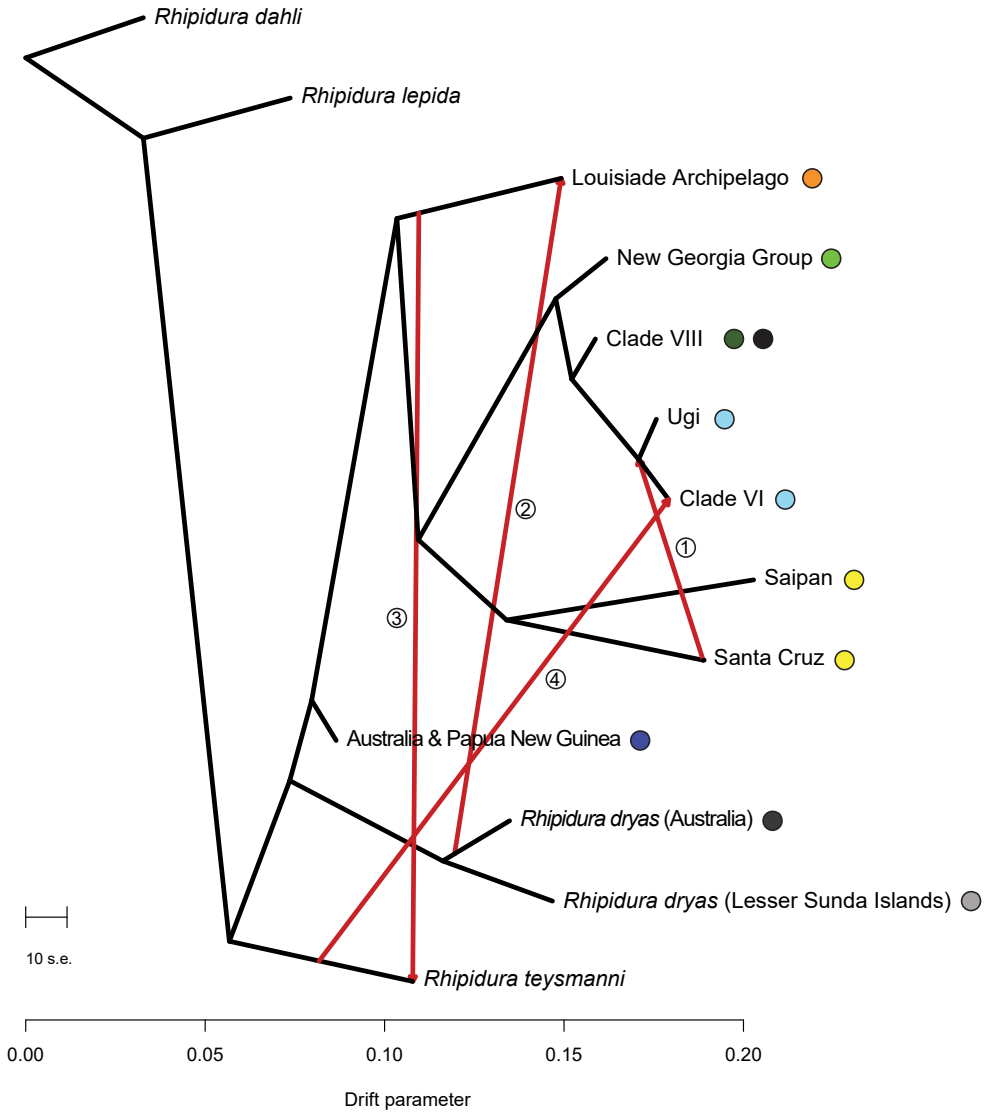


Figure 3. Species tree estimated in TreeMix for the *Rhipidura rufifrons* species complex using the 70% minimum representation dataset of SNPs (5625 SNPs). Migration edges are numbered in the order that they were added and explain 0.7302%, 0.4659%, 0.3452%, and 0.3995% of the variation in the SNP data, respectively. Tips are labeled with geographic locations for *R. rufifrons* samples and the RAxML (Figure 1) clade assignments.

TreeMix analyses the New Georgia Group (Clade VI; *R. r. granti*) was recovered as sister to Clades V and VII (Figure 3). With no migration edges, the species tree explained 97.58% of the variation in the SNP data. We added migration edges until they explained less than 0.2% of the data, resulting in four migration events. The first migration edge indicated gene flow between the Santa Cruz population and the ancestor of the populations on Makira (*R. r. russata*), Ugi (*R. r. ugiensis*), and Santa Catalina (*R. r. kuperi*; Figure 3). The second migration edge paired the Australian populations of *R. dryas* and the Louisiade Archipelago (*R. r. louisiadensis*) population. The third migration edge again involved the Louisiade Archipelago population, but this time *R. r. louisiadensis* was hypothesized to undergo gene flow with *R. teysmanni* (Sulawesi). The last migration edge explained 0.3995% of the variation and indicated potential introgression between *R. teysmanni* (Sulawesi) and Clade V (Makira and Santa Catalina).

## Discussion

Current taxonomy of the *R. rufifrons* species complex is based on geography and qualitative descriptions of plumage and vocalization differences, resulting in 18 described subspecies (Clements et al., 2019). Here, we examined relationships among 12 *R. rufifrons* subspecies in an explicit phylogenetic context and found variable amounts of genomic divergence across the complex. Below we discuss the biogeographic importance of the Louisiade Archipelago, diversification for island isolates, and phylogeography of the Solomon Islands in light of other species complexes from the same region.

### Louisiade Archipelago

The Louisiade Archipelago, located southeast of Papua New Guinea, harbors a diverse set of endemic taxa (Allison & Leisz, 2009; Linck, Schaack, & Dumbacher, 2016; Polhemus, Englund, & Allen, 2004), and recent phylogenetic studies (e.g., Andersen et al., 2014; Andersen

et al., 2015; Kearns, Joseph, & Cook, 2013; Oliver, Rittmeyer, Kraus, Richards, & Austin, 2013; Pedersen et al., 2018; Tu, Yang, Liang, & Zhang, 2018) have supported the independence of endemic lineages with genetic data. In the most dramatic examples (e.g., Andersen et al. 2014), the Louisiade populations are highly divergent and sister to species complexes that span the Australasian region. Similar to these studies, we recovered a deep phylogenetic split between *R. r. louisiadensis* and all other subspecies of *R. rufifrons* (Figure 1). However, unlike Andersen et al. (2014), we recovered a pattern more consistent with the Louisiade Archipelago acting as an early stepping stone in the colonization of Melanesian and Micronesian islands. Regardless, the growing body of evidence across taxonomic groups (Andersen et al., 2014; M J Andersen et al., 2015; Oliver, Travers, Richmond, Pikacha, & Fisher, 2017; Shashank, Chakravarthy, Raju, & Bhanu, 2014) indicates a potentially important role for the Louisiade Archipelago in the early diversification of lineages across the Southwest Pacific.

The geographically proximate island of New Guinea has played a significant role in diversification within the genus *Rhipidura*. New Guinea contains both highland and lowland species from throughout the phylogeny for the *Rhipidura* genus and New Guinea holds more species than any other geographic location (Nyári et al., 2009). However, New Guinea is conspicuously absent from the distribution of the otherwise widespread *R. rufifrons* complex. The absence of *R. rufifrons* on New Guinea could be a consequence of many closely related species having already diversified on the island, in particular *R. rufidorsa* and *R. brachyrhycha* which are members of the same *Rhipidura* subclade as *R. rufifrons* (Nyári et al., 2009).

**Island Isolates**

The two most geographically remote taxa in our dataset, *R. r. saipanensis* (Northern Mariana Islands) and *R. r. agilis* (Santa Cruz) are separated from the nearest sampled *R. rufifrons*

population by over 2,600 km and 400 km, respectively. Furthermore, they are separated from one another by a distance of over 3,600 km of open ocean, and yet were recovered as sister taxa in our analysis. The perhaps surprising sister relationship of these lineages suggests that a single dispersive ancestor may have quickly colonized archipelagos across the Pacific Ocean. Unfortunately, tissue samples were unavailable (or, in the case of *R. r. mariae*, the generation of sequence data was unsuccessful) for the following taxa: two subspecies from the Caroline Islands (*R. r. versicolor* [Yap] and *R. r. kubaryi* [Pohnpei]), one extant subspecies and one extinct subspecies from the Mariana Islands (*R. r. mariae* [Rota and Agiguan] and *R. r. uraniae* [Guam], respectively) and two subspecies from the Santa Cruz Islands (*R. r. melaenolaema* [Vanikoro] and *R. r. utupuae* [Utupuae]). Thus, it is difficult to determine the relative importance of multiple colonization events or *in situ* diversification in the evolutionary history of the *R. rufifrons* complex. In other avian systems with isolated island archipelago populations, researchers have shown that co-occurring lineages often are not sister taxa (Cibois et al., 2011; Cibois, Thibault, & Pasquet, 2007; Ryan, Klicka, Barker, & Burns, 2013). Further, it is becoming increasingly clear that avian lineages on the Mariana Islands frequently have unexpected evolutionary relationships. For example, Cibois et al. (2011) and Andersen et al. (2015) both found that populations on the Mariana Islands were not monophyletic, but concordant patterns were mostly absent between the two systems. Here, we find support for the association of the far-flung Mariana Islands with another remote island population, but a more complete sampling of Pacific lineages would enable a more robust reconstruction of the biogeographic history of the *R. rufifrons* complex.

**Solomon Islands**

Currently seven taxa are recognized within the *R. rufifrons* species complex in the Solomon Islands: *R. r. ugiensis*, *R. r. russata*, *R. r. kuperi*, *R. r. granti*, *R. r. commoda*, *R. r. rufofronta*, and *R. r. brunnea*. However, we found genetic evidence for only three genomic backgrounds (Figure 1d) with several admixed individuals. Similarly, phylogenetic analyses divided the Solomon Island samples into three major clades (Clade V–VII), but also identified substructure within Clades V and VII not indicated in clustering analyses. For example, within Clade VII, samples from Bougainville and the Shortland Islands formed a well-supported clade, as did individuals from Malaita and Guadalcanal. However, little genetic structure existed between populations from Malaita (*R. r. brunnea*) and Guadalcanal (*R. r. rufofronta*). Although Malaita is an oceanic island that hosts many endemic bird species (Mayr & Diamond, 2001), we did not find evidence for a distinct Malaitan taxon. *Rhipidura rufifrons* may have recently colonized Malaita without sufficient time for identifiable genetic differentiation. With recent colonization, we would expect Malaitan populations to exhibit lower genetic diversity compared to other islands, contrary to our results (Figure 2). Alternatively, recent or ongoing gene flow between Guadalcanal and Malaita may have prevented genetic divergence between the two populations while maintaining relatively high genetic diversity. STRUCTURE analyses did not contradict the hypothesis of ongoing gene flow between these populations. In fact, STRUCTURE suggested a broader pattern of isolation by distance (see Gene Flow) that not only included all the populations from Clade VII (Figure 1d) but also the New Georgia Group individuals that form Clade VI. Glacial cycling in the Pleistocene would have created many land bridge connections within the Solomon Islands (Mayr & Diamond, 2001) and are likely a cause for the close association between avifauna on some islands. However, the pattern for *R. rufifrons* cannot be explained simply by Pleistocene glacial cycles because although Bougainville,

Choiseul, and Isabel were most likely connected by land bridges to form the "Greater Bukida Islands", it is unclear if Guadalcanal was ever physically connected with this group (Mayr & Diamond, 2001). Furthermore, Malaita and the New Georgia Group are surrounded by deep water and certainly did not form land bridge connections with any other islands.

Populations from the island of Makira (*R. r. russata*) and its satellites Ugi (*R. r. ugiensis*) and Santa Catalina (*R. r. russata*) in the southeastern part of the Solomon Archipelago form another well-supported clade. Previous taxonomic hypotheses based on plumage variation are corroborated by genomic data presented here, recognizing *R. r. ugiensis* as a distinct taxon from populations on Makira (*R. r. russata*) and Santa Catalina (*R. r. kuperi*). *Rhipidura r. ugiensis* is an isolated taxon on a small island with distinct melanistic plumage along the throat and chin, which mirrors several well-documented examples of the evolution of melanism on small islands in this region (Uy et al., 2016; Uy, Moyle, Filardi, & Cheviron, 2009; Uy & Vargas-Castro, 2015). Although *R. r. ugiensis* individuals were well-supported as an independent lineage, the Makira and Santa Catalina populations were not recovered as reciprocally monophyletic. Instead, Santa Catalina individuals formed a clade embedded within a paraphyletic group from Makira (Figure 3, Clade V). These results suggest that *R. r. kuperi* is a result of recent colonization of Santa Catalina by the Makira population, with insufficient time for complete lineage sorting.

**Gene Flow**

Although we recovered phylogenetic structure among many closely related allopatric populations, we also found evidence of possible gene flow between allopatric taxa across vast geographic distances. For example, TreeMix indicated gene flow between *R. r. agilis* (Santa Cruz, Clade III) and the common ancestor of Clade V (Ugi, Makira and Santa Ana). The migration edge between *R. r. agilis* (Santa Cruz) and Clade V accounted for 0.73% of the

variance in the genomic dataset. In addition to TreeMix analyses, STRUCTURE analyses for all values of K indicated a shared genomic background for Santa Cruz individuals and individuals from Ugi, Makira and Santa Catalina. Therefore, despite nearly 400 km of open ocean between them, these populations have maintained shared genomic variation.

The placement of the first migration edge discussed above was within a region of the phylogeny with rather robust sampling, and thus we are confident in its reality. However, the interpretation of the other three inferred migration edges is less straightforward because the three additional migration edges involved outgroup taxa and the placement of these migration edges would likely be influenced by the inclusion of potentially independent ingroup lineages such as *R. r. torrida* (Molucca Islands) or *R. rufifrons* individuals from Rossel Island (Louisiade Archipelago). For example, Linck et al. (2016) revealed the Rossel population of *Zosterops griseotinctus* as a distinct lineage, and demostrated that the Rossel Island population did not form a clade with individuals from the remainder of the Louisiade Archipelago islands. Unfortunately, modern sampling was not available for Indonesian taxa like *R. r. torrida*, and the sample from Rossel Island we were able to acquire did not produce enough useable data to be included in our final dataset. Thus, while we find support for gene flow between Santa Cruz and Clade V, we caution against over-interpretation for the other three migration events inferred by TreeMix until more complete sampling for *R. dryas* and *R. rufifrons* is available.

Although individuals from Bougainville, Shortland Islands, Choiseul, and Isabel (Clade VII) shared 30% or more of their genomic background with New Georgia Group (Clade VI) samples in a K = 3 population model for Solomon Island samples (Figure 1d), none of the four migration events inferred by TreeMix involved either of these clades. Phylogenetic analyses recovered the monophyly of Clade VI and the monophyly of the Bougainville and Shortland

Island individuals with strong support, but both groups contained very few fixed differences (Figure 2). Considering the contrasting results recovered by clustering, phylogeny, and species tree analyses, populations from the northern Solomon Islands clearly warrant further investigation.

## Conclusions

In this study, we identified several biogeographic patterns of broad interest. First, the Louisiade Archipelago contains a distinct population of the *R. rufifrons* complex that shares an old relationship with the rest of the Pacific lineages. In combination with similar patterns in other taxonomic groups, our results indicate that this small set of islands may have been important for the early diversification of *R. rufifrons* and other terrestrial lineages. We also discovered that individuals from of the Northern Mariana Islands (*R. r. saipanensis*) formed a sister relationship with Santa Cruz individuals, again highlighting a common pattern whereby birds from the Mariana Islands do not form sister relationships to the nearest sampled conspecific population (Andersen et al., 2015; Cibois et al., 2011). Within the Solomon Islands: Malaita, the New Georgia Group, and Makira lacked Pleistocene land bridges to the other major Solomon Islands, but in considering individuals from these islands, only *R. r. brunnea* from Malaita was not supported as an independent lineage. This research adds to the growing body of literature for systems exhibiting rapid phenotypic evolution despite recent or ongoing gene flow in some situations. Furthermore, despite varied life histories and dispersal ability between taxa representing rapid radiations, convergent patterns of diversification in the Indo-Pacific are beginning to emerge.

# Chapter 2: Where high dispersal meets high diversification: The case of the White-bellied cuckooshrike, *Coracina papuensis*

## Introduction

Gene flow is generally regarded as a homogenizing force that will inhibit population differentiation if not interrupted and thus, in theory, an inverse relationship should exist between gene flow and genetic divergence (Wright, 1931). Dispersal ability is a common proxy for gene flow, and although challenges exist in estimating dispersal ability (Claramunt, Derryberry, Remsen, & Brumfield, 2012; Weeks & Claramunt, 2014), recent work, with few exceptions (Owens, Bennett, & Harvey, 1999; Phillimore, Freckleton, Orme, & Owens, 2006; Weeks & Claramunt, 2014), has supported the theorized inverse relationship between dispersal ability and differentiation (Belliure, Sorci, Møller, & Clobert, 2000; Bolmgren & Eriksson, 2005; Claramunt et al., 2012; Harvey et al., 2017; Kennedy et al., 2016; Riginos, Buckley, Blomberg, & Treml, 2014; Smith et al., 2014). In birds, dispersal ability is often quantified by wing morphology or an ecological characteristic such as foraging stratum, whereby canopy foraging birds would be classified as high dispersers and understory species characterized as low dispersers. Smith et al. (2014) used foraging stratum as an indicator of dispersal ability for species of Neotropical birds and found a significant increase in species diversity for understory foragers relative to canopy species, consistent with theoretical expectations (Wright, 1931). They concluded that diversification in the Neotropics was mostly independent of vicariant landscape changes like mountain orogeny or river formation. Smith et al. (2014) instead suggested that dispersal across barriers an already formed landscape was a major driver of speciation in the Neotropics. In contrast to continental geographies, island systems like those in the southwest Pacific present more discontinuous landscapes that often required dispersal between isolated suitable habitats, and it is unknown if patterns of differentiation would mirror those in relatively continuous habitats of continental systems. In fact, Melanesian archipelagos inspired the “intermediate-

dispersal" hypothesis (Diamond et al., 1976; Mayr & Diamond, 2001) which posits that diversification should be maximized at a dispersal ability high enough to enable colonization of novel habitats but insufficient to homogenize isolated populations.

Within the Indo-Pacific, understory bird species have been the focus of much phylogeographic research (Andersen, Hosner, Filardi, & Moyle, 2015; Andersen et al., 2013; Campillo, Manthey, Thomson, Hosner, & Moyle, 2019; Campillo, Oliveros, Sheldon, & Moyle, 2018; Fabre et al., 2013; Filardi & Moyle, 2005; Irestedt et al., 2013; Manthey et al., 2017; Nyári et al., 2009), whereas canopy foragers in the region remain largely unstudied using similar approaches, a bias most likely ascribed to increased difficulty in obtaining sufficient sampling of canopy species. For understory species, taxonomic classifications often failed to accurately reflect diversity. In some cases, genetic analyses identified additional taxa or non-monophyletic groups (Andersen et al., 2015; Andersen et al., 2013; Andersen et al., 2015), whereas other investigations provided evidence (Filardi & Moyle, 2005) that challenged early assumptions for a one way pattern of colonization from continents to islands (Mayr, 1941). Thus, phylogeography studies of understory species have informed taxonomy and, in some cases, even reshaped the understanding for how islands are important in diversification. Thus, similar genetic studies of canopy species are needed to allow comparison of phylogeographic structure between understory and canopy species in island systems, and between island and continental systems.

The White-bellied Cuckooshrike (*Coracina papuensis*) is distributed across the Indo-Pacific: Australia (4 taxa), Papua New Guinea (6 taxa; including Manus Island, Louisiade Archipelago, and Bismarck Archipelago), Solomon Island Archipelago (3 taxa), and western Indonesia (2 taxa). *Coracina papuensis* occurs sea level to elevations of 800m in rainforests, secondary forest, human disturbed areas (i.e., gardens, farmland, teak and pine plantations),

savannahs and riparian areas. They are most commonly observed foraging on large insects in the forest canopy either alone or in small groups (Taylor, 2019). Morphological differences between subspecies of *C. papuensis* are subtle, and include variation in the facial mask, darkness of plumage for head and upper parts, coloration on primary feather edges, ventral bars on females, and body size. Currently, 13–15 subspecies have been described that categorizes the diversity within the *C. papuensis* species complex. Given that previous work suggests a lower diversification rate for canopy species (Smith et al., 2014), *C. papuensis* may be either the exception to the rule or a group in which current taxonomy does not reflect its true diversity. Thus, the White-bellied cuckooshrike (*Coracina papuensis*) presents an ideal canopy-foraging species complex to add to an expanding body of southwest Pacific phylogeographic literature.

The distribution of *C. papuensis* overlaps broadly with several previously examined understory lineages, particularly within the Solomon Island archipelago, permitting direct comparisons of diversification between understory and canopy-foraging species. Furthermore, the 13–15 taxa within *C. papuensis* is on par with other complexes which are well known for rapid diversification and are referred to as the "great-speciators" (Mayr & Diamond, 2001). However, a phylogenetic analysis of *C. papuensis* has yet to be completed and thus the taxonomy for subspecies within the *C. papuensis* group may not reflect the evolutionary history of the lineage (Zink, 2004). Therefore, this study aims to reconstruct the evolutionary history of *C. papuensis*, identify independent lineages within the group, place diversification of *C. papunesis* in context by comparing it to species with low dispersal ability with which it is co-distributed, and investigate the influence of gene flow on diversification across continental and island landscapes.

## Methods

**Sampling**

Fresh tissues were sampled from 49 individuals of *Coracina papuensis* (Table 2, Figure 4) representing 10 of the 13 described subspecies (Clements et al., 2019) and a representative for the dark morph of *C. p. robusta*. Two individuals of *Coracina caledonica* were used as an outgroup for analyses (Jønsson et al., 2010). We extracted genomic DNA from all individuals using a QIAGEN DNeasy blood and tissue kit and quantified DNA concentrations of the extracts with a Qubit Flourometric Quantitation (Life Technologies).

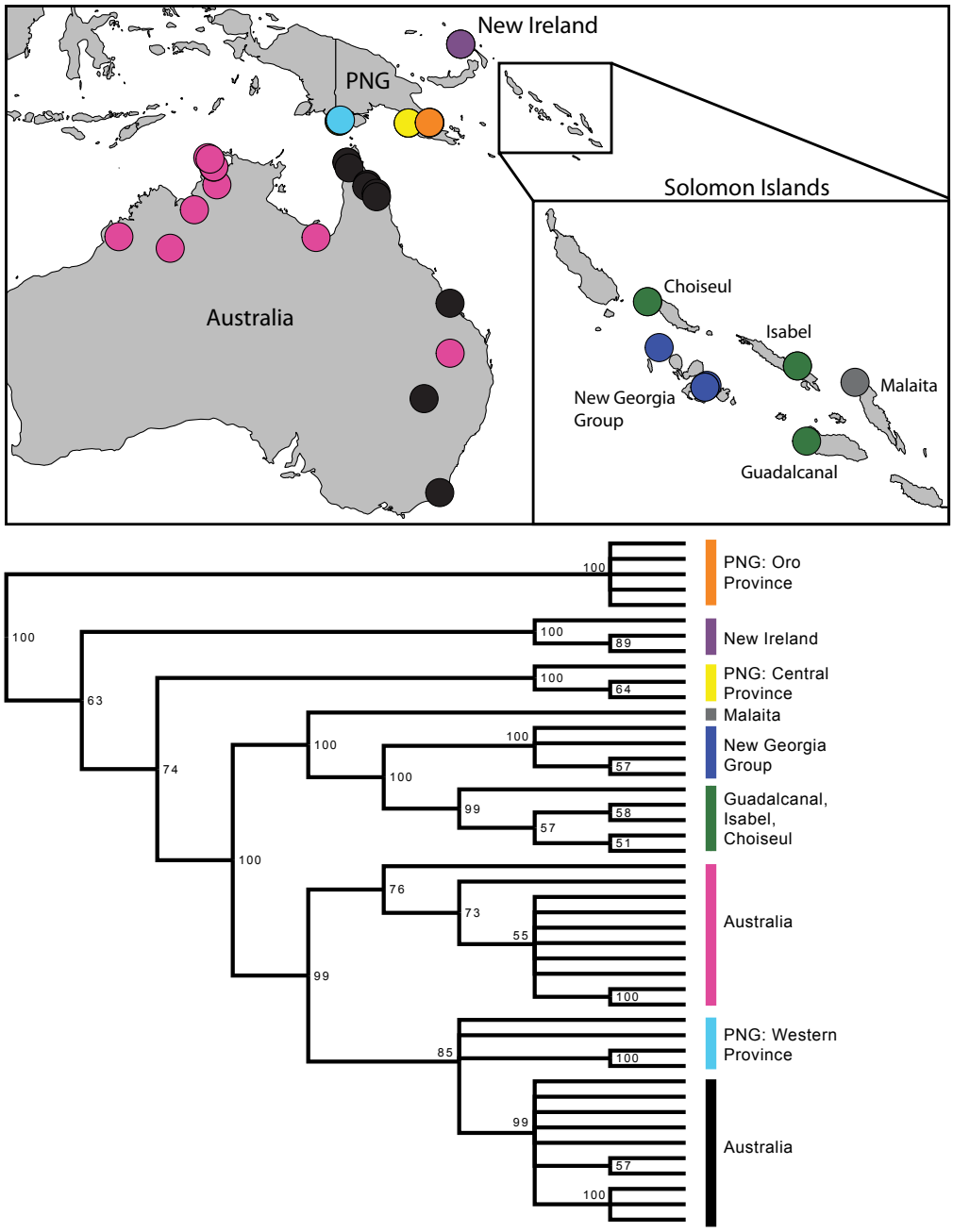


Figure 4. Sampling locations for *C. papuensis* are indicated with circles on the map of the Australo-Papuan region. Colored circles correspond to the colored bars on the phylogeny. The species tree was obtained by analyzing a single SNP per locus for the 50% minimum representation dataset (5636 SNPs) with SVDquartets. Nodes receiving less that 50% bootstrap support were collapsed. The outgroup has been removed from the figure.

Table 2. Locality and voucher information for *Coracina* specimens included in chapter 2

| Species | ID | Museum | Country | Lat | Long | Cov Median |
|---|---|---|---|---|---|---|
| *C. papuensis apsleyi* | B50598 | ANWC | AUS | -17.67 | 123.57 | 31 |
| *C. papuensis apsleyi* | B55352 | ANWC | AUS | -18.55 | 127.69 | 19 |
| *C. papuensis artamoides* | B29427 | ANWC | AUS | -17.73 | 139.39 | 22 |
| *C. papuensis artamoides* | B29749 | ANWC | AUS | -11.99 | 141.88 | 19 |
| *C. papuensis artamoides* | B29750 | ANWC | AUS | -11.99 | 141.88 | 25 |
| *C. papuensis artamoides* | B32715 | ANWC | AUS | -12.43 | 142.04 | 18 |
| *C. papuensis artamoides* | B42982 | ANWC | AUS | -13.68 | 143.51 | 23 |
| *C. papuensis artamoides* | B43075 | ANWC | AUS | -13.84 | 143.46 | 29 |
| *C. papuensis artamoides* | B43532 | ANWC | AUS | -22.74 | 150.14 | 15 |
| *C. papuensis artamoides* | B51428 | ANWC | AUS | -14.36 | 144.22 | 16 |
| *C. papuensis artamoides* | B51487 | ANWC | AUS | -14.62 | 144.25 | 23 |
| *C. papuensis hypoleuca* | 60701 | UWBM | AUS | -13.72 | 131.43 | 11 |
| *C. papuensis hypoleuca* | B33463 | ANWC | AUS | -12.38 | 131.17 | 10 |
| *C. papuensis hypoleuca* | B33648 | ANWC | AUS | -12.35 | 131.20 | 16 |
| *C. papuensis hypoleuca* | B48579 | ANWC | AUS | -11.65 | 130.70 | 12 |
| *C. papuensis hypoleuca* | B48695 | ANWC | AUS | -11.76 | 130.88 | 29 |
| *C. papuensis hypoleuca* | B51067 | ANWC | AUS | -15.62 | 129.63 | 9 |
| *C. papuensis robusta* | B41369 | ANWC | AUS | -26.55 | 150.15 | 16 |
| *C. papuensis robusta* | B44854 | ANWC | AUS | -37.17 | 149.32 | 13 |
| *C. papuensis robusta* | B49305 | ANWC | AUS | -30.01 | 148.06 | 19 |
| *C. papuensis angustifrons* | B55902 | ANWC | PNG | -9.01 | 146.81 | 12 |
| *C. papuensis angustifrons* | B55926 | ANWC | PNG | -9.00 | 146.80 | 26 |
| *C. papuensis angustifrons* | B55957 | ANWC | PNG | -9.02 | 146.80 | 24 |
| *C. papuensis angustifrons* | B56028 | ANWC | PNG | -9.00 | 146.79 | 24 |
| *C. papuensis angustifrons* | B57661 | ANWC | PNG | -8.99 | 148.52 | 8 |
| *C. papuensis angustifrons* | B57684 | ANWC | PNG | -8.98 | 148.52 | 11 |
| *C. papuensis angustifrons* | B57737 | ANWC | PNG | -8.95 | 148.52 | 13 |
| *C. papuensis angustifrons* | B57759 | ANWC | PNG | -8.97 | 148.48 | 10 |
| *C. papuensis angustifrons* | B57868 | ANWC | PNG | -8.97 | 148.48 | 10 |
| *C. papuensis oriomo* | B56121 | ANWC | PNG | -8.86 | 141.26 | 11 |
| *C. papuensis oriomo* | B56157 | ANWC | PNG | -8.87 | 141.24 | 22 |
| *C. papuensis oriomo* | B56226 | ANWC | PNG | -8.82 | 141.30 | 21 |
| *C. papuensis oriomo* | B56232 | ANWC | PNG | -8.78 | 141.34 | 16 |
| *C. papuensis sclaterii* | 27758 | KUNHM | PNG | -3.00 | 151.00 | 25 |

| Species | ID | Museum | Country | Lat | Long | Cov Median |
|---|---|---|---|---|---|---|
| *C. papuensis sclaterii* | 27769 | KUNHM | PNG | -3.00 | 151.00 | 20 |
| *C. papuensis sclaterii* | 27776 | KUNHM | PNG | -3.00 | 151.00 | 26 |
| *C. papuensis elegans* | 33896 | KUNHM | SI | -7.72 | 156.75 | 17 |
| *C. papuensis elegans* | 60256 | UWBM | SI | -9.48 | 159.98 | 13 |
| *C. papuensis elegans* | 60356 | UWBM | SI | -9.54 | 159.64 | 14 |
| *C. papuensis elegans* | 63014 | UWBM | SI | -8.46 | 157.69 | 13 |
| *C. papuensis elegans* | 66005 | UWBM | SI | -8.49 | 157.65 | 9 |
| *C. papuensis elegans* | 66049 | UWBM | SI | -8.49 | 157.65 | 28 |
| *C. papuensis eyerdami* | 19424 | KUNHM | SI | -8.40 | 160.59 | 22 |
| *C. papuensis perpallida* | 58716 | UWBM | SI | -8.08 | 159.46 | 13 |
| *C. papuensis perpallida* | 58717 | UWBM | SI | -8.08 | 159.46 | 21 |
| *C. papuensis perpallida* | 63076 | UWBM | SI | -6.82 | 156.53 | 27 |
| *C. papuensis perpallida* | 63194 | UWBM | SI | -6.83 | 156.52 | 14 |
| *C. caledonica* | 60241 | | | | | 17 |
| *C. caledonica* | 60281 | | | | | 12 |

Coverage median refers to the median depth of coverage per locus. AUS Australia, PNG Papua New Guinea, SI Solomon Islands. ANWC Australian National Wildlife Collection, KUNHM University of Kansas Natural History Museum, UWBM University of Washington Burke Museum

**Sequencing**

We obtained numerous putatively neutral anonymous genetic loci by performing a modified single digest RAD-seq protocol (Miller et al., 2007). We used NdeI to digest genomic DNA and then ligated custom adapter sequences with individual barcodes (Andolfatto et al., 2011). We pooled samples and then selected for a fragment size range between 450-600 bp using a Pippin Prep electrophoresis cassette (Sage Science). The pooled library was examined for DNA quality and quantity following a brief PCR (98 C for 30 s, 14 cycles of 98 C for 10 s, 64 C for 30 s, and 72 C for 20 s with a final extension of 72 C for 7 min). Sequencing was then conducted at the University of Kansas Genome Sequencing Core Facility using an Illumina HiSeq 2500 for 100 bp single-end reads.

**Bioinformatics**

We used the STACKS v 2.41 (Catchen et al., 2011) pipeline to assemble loci from the fastQ output acquired from the Illumina sequencing run. Using the *process_RADtags* script from STACKS, individuals were de-multiplexed and low-quality reads were removed using default settings. We used the *ustacks* module to ascertain loci within individuals. In order to form a set of putative loci, identical reads were first formed into stacks if they met the minimum depth of coverage (-m) which we varied across analyses (m = 3, 5, 7). Different stacks within individuals were then combined into potential loci if they contained fewer differences than permitted by the -M parameter and we examined datasets using -M values ranging from 1–5. Next, we ran the *cstacks* module to combine individual loci into a catalog of loci. We also examined a range of values for number of mismatches permitted between individuals (n = 1, 3, 5). We then matched the sequence data for each individual to the catalog using the default settings of the *sstacks* module, transposed the dataset with *tsv2bam*, and then built consensus sequences for loci with *gstacks*. Datasets were constructed with the *populations* module within STACKS. We required loci to have 50% representation across all individuals (-r 0.5), and then filtered for minor allele frequency (> 0.05) and observed heterozygosity (< 0.5).

**Phylogenetics and Population Genetics**

Prior to phylogenetic analyses, each single nucleotide polymorphism (SNP) was collapsed into a consensus allele for each individual, and we identified polymorphic SNPs with ambiguity codes. We concatenated the consensus alleles for each individual and retained a dataset that included all available SNPs in addition to a dataset that was limited to a single SNP per locus. We inferred phylogenetic relationships among individuals in a maximum likelihood (ML) framework with RAxML v8 (Stamatakis, 2014) using 1000 rapid bootstraps and a GTR + GAMMA model of sequence evolution. In situations where the most common gene tree is in

conflict with the true species tree (i.e. the anomaly zone: Degnan & Rosenberg, 2006; Degnan & Salter, 2005) concatenation will fail to recover accurate phylogenetic relationships. Therefore, we also inferred species tree topology using SVDquartets (Chifman & Kubatko, 2014) in PAUP* (Swofford, 1998). We sampled all possible quartets, used a single SNP per locus, and completed 500 bootstrap replicates. Individual quartets were amalgamated into a species tree within PAUP* using a *Quartet FM* algorithm (Reaz, Bayzid, & Rahman, 2014)

The outgroup *C. caledonica* was removed for analyses using the genetic clustering programs STRUCTRE (Pritchard et al., 2000) and Discriminant Analysis of Principle Components (DAPC; Jombart et al., 2010). We retained a single SNP per locus to reduce the influence of linkage on the analyses. We completed ten independent structure analyses for a range of K values (1–10). We ran each analysis for 550,000 MCMC generations with the first 50,000 generations used as burn-in. We assessed the likelihood scores for different K values and a **Δ**K (Evanno et al., 2005) calculation to estimate the number of genetic clusters present in the dataset. STRUCTURE assumes no hierarchal population relationships and will minimize deviation from Hardy-Weinberg equilibrium within each genetic cluster, therefore we also used a multivariate approach (DAPC) to identify genetic clustering. For DAPC, analyses were completed in R (R Core Development Team, 2012) with the package 'adegenet' (Jombart, 2008; Jombart & Ahmed, 2011), and we estimated the most likely number of populations using the Bayesian Information Criterion (BIC). We addressed potential over-parameterization by estimating the optimal number principle components (PC) for the data using 'optim.a.score'. Analyses were initially completed with the estimated optimal number of PCs, and to determine consistency of results, the number of PCs permitted was varied in subsequent analyses.

## Results

**SNP Data**

We retained a total of 68,821,754 reads from 49 samples after removal of low-quality reads. Reads per individual ranged from 4,104,032–148,700 (median = 1.087 million). We did not identify obvious differences in population differentiation ($F_{st}$) or population assignment (DAPC) between the data matrices produced by varying parameters within the STACKs pipeline. Therefore, we only present results from analyses of the data matrix produced using the STACKs parameters of -M 2, -m 3, and -n 3. We eliminated loci not present in 50% of individuals (-r 0.5) and the resulting dataset was 84.1% complete, with 5,636 RAD loci, containing a total of 8,633 SNPs.

**Population Genetics**

DAPC analyses favored a four-population model using the BIC. The four-population model grouped individuals from the following locations into distinct clusters: New Ireland (1), Oro and Central Province of Papua New Guinea (PNG) (2), Solomon Islands (3), and Australia with the Trans-fly region of PNG (4). Although DAPC assigned Oro and Central Province individuals to the same genetic cluster, the samples were otherwise grouped consistent with their collection localities (Figure 5). Two individuals, however, grouped with clusters not of their geographic origin: one individual from Central Province of PNG (ANWC B55902) grouped with the Solomon Islands and one individual from Guadalcanal (UWBM 60256) was grouped with Australia (Figure 5).

An evaluation of ΔK and likelihood scores for independent STRUCTURE runs from 1–10 identified K = 3 as the most likely population model. One cluster included individuals from New Ireland and Oro Province, a second united individuals from the Solomon Islands, and a third cluster grouped Australian and Trans-fly samples (Figure 5). All Central Province

individuals were of an admixed genetic background (Figure 5). One individual from Central Province (ANWC B55902) was assigned to all three genetic clusters, but the largest proportion of the genomic background was more consistent with that of an individual from the Solomon Islands (Figure 5). In K = 3 analyses, two individuals had genomic backgrounds divergent from those of geographically proximate individuals. A sample from Guadalcanal (KU 60256) shared a majority of its genomic background (~75%) with Australian samples.

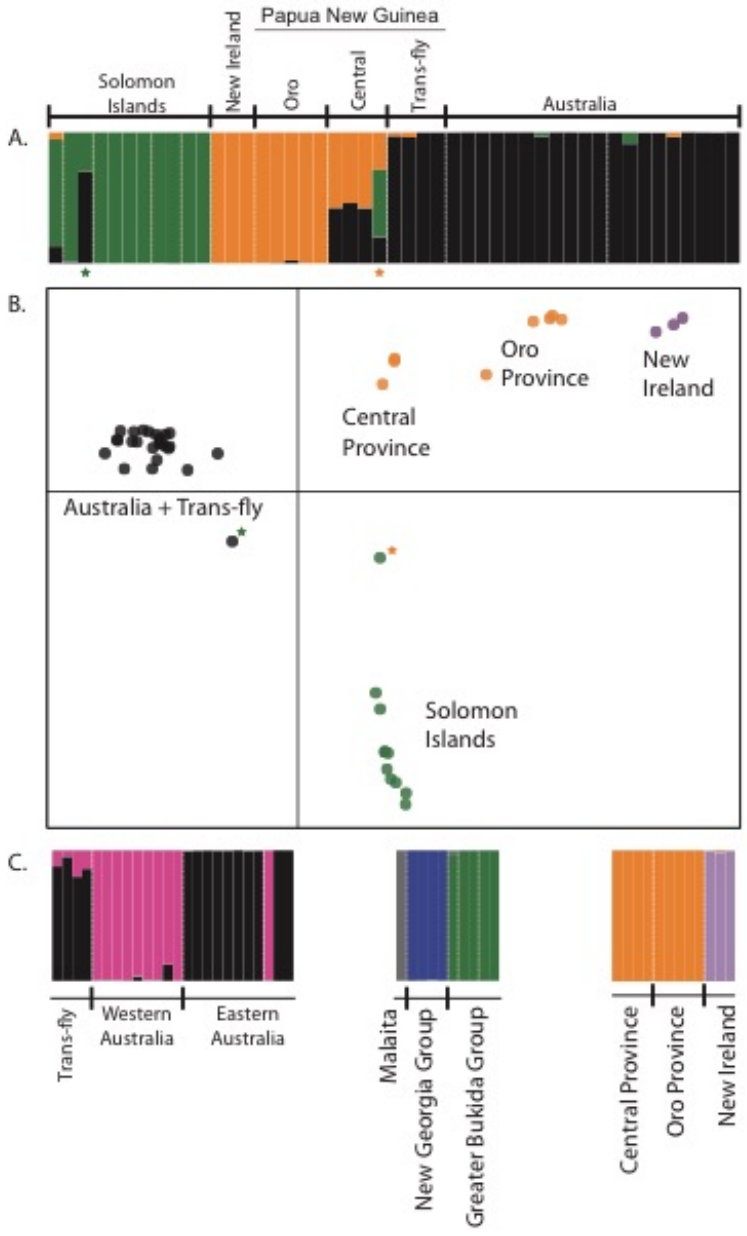


Figure 5. A) The preferred population model (K = 3) from STRUCTURE when using all ingroup samples and a single SNP per locus. B) DAPC analyses on the same single SNP per locus dataset. The orange star indicates the sample ANWC B55902 from Central Province and the green star indicates the sample UWBM 60256 from Guadalcanal. C) The preferred population models after reanalyzing the single SNP per locus dataset subset into Australian + Trans-fly (K = 2), Solomon Island (K = 3), and Central Province + Oro Province + New Ireland (K = 2) individuals.

**Phylogeny**

The two individuals identified in clustering analyses as having significant admixture (KU 60256 and ANWC B55902) were ultimately excluded from phylogenetic analyses. SVDquartets analyses for the single SNP per locus dataset produced a phylogeny with generally well-supported nodes, with few exceptions. Individuals from Oro Province (PNG), New Ireland, Central Province (PNG), and the Solomon Islands were each recovered as clades with 100% bootstrap (BS) support (Figure 4). Within the Solomon Islands, individuals from Guadalcanal, Choiseul, and Isabel formed a clade (BS = 99), New Georgia Group individuals were united as a clade (BS = 100), and the single Malaitan individual was sister to all other individuals from the Solomon Islands (Figure 4). Individuals sampled from Australia and the Western Province of PNG (i.e., Trans-fly region) formed a well-supported clade (BS =99). With a single exception, individuals from Western Australia formed a clade (BS = 76) and individuals from eastern Australia (i.e., Queensland and New South Wales) were united with individuals from the Western Province of PNG (i.e., Trans-fly region) with BS = 85. The lone exception to the east-west division (ANWC B41369) was from southeastern Queensland and was embedded within the clade of Western Australia samples (Figure 4). Strong support was recovered for ingroup monophyly, but relatively low support (BS = 63) was recovered for Oro Province individuals as sister to all other *C. papuensis* populations (Figure 4). We found higher support (BS=74) for a clade containing individuals from Central Province (PNG), Solomon Islands, Western Province (PNG), and Australia. Results from RAxML analyses were mostly concordant with few differences. RAxML analyses recovered the same pattern with respect to individuals from Australian and the Trans-fly region, but with stronger support for reciprocally monophyly (BS = 100 and BS = 99). Further, lower support was indicated (BS = 56) for Queensland and New

South Wales as a clade to the exclusion of the Trans-fly region of PNG. Unlike SVDquartets, RAxML did not support Central Province individuals as a clade.

## Discussion

Phylogenetic and phylogeographic analysis of genomic data revealed substantial population structure within *C. papuensis*. Overall, we established genomic support for distinct populations throughout the *C. papuensis* complex, but fewer than indicated by taxonomy. We also found evidence that several individuals are the result of recent admixture between differentiated populations, some of which are separated by hundreds of kilometers of open ocean. Below we examine the diversification of *C. papunesis* across the Sahul shelf (i.e., Australia and New Guinea), Solomon Island Archipelago, and discuss the role of dispersal on diversification for *C. papuensis*.

### Australo-Papuan Phylogeography

The populations of *C. papuensis* distributed across Australia and New Guinea have been divided into as many as eight currently recognized taxa: *C. p. papuensis* (NG), *C. p. angustifrons* (NG), *C. p. intermedia* (NG), *C. p. oriomo* (NG), *C. p. apsleyi* (AU), *C. p. hypoleuca* (AU), *C. p. artamoides* (AU), and *C. p. robusta* (AU) (Clements et al., 2019; Taylor, 2019). Despite this previously recognized diversity and a distinctive dark plumage morph (*C. p. robusta*), analyses supported, at most, two Australian lineages. Furthermore, the geographically isolated individuals from the Trans-fly region of New Guinea were embedded within the clade of individuals from Australia. *Coracina papunesis* is a habitat generalist and can be found in woodlands, riparian areas, rainforest, forest edges, gardens, mangroves and plantations. Factoring in the ecology of *C. papuensis* as a canopy forager and therefore a presumed high dispersal species, it is unsurprising that we observed limited differentiation for the group across Australia.

Though we did not find strongly-supported sub structure within Australia, the two sub-clades that received some phylogenetic support generally followed an east-west division. Similar to previous studies (Peñalba, Joseph, & Moritz, 2019; Pepper et al., 2017) but in contrast to others (Toon, Drew, Mason, Hughes, & Joseph, 2017), we found individuals from the Trans-fly region grouping with individuals from eastern Australia. STRUCTURE analysis using only Australian and Trans-fly individuals supported a two-population "east-west" model with only one south-east Queensland sample breaking the trend (ANWC B41369), but STRUCTURE results also indicated some admixture for Trans-fly individuals. The potential for ongoing gene flow, shared ancestral polymorphism or recent diversification between western Australia and Trans-fly individuals has also been observed in other avian and non-avian systems (Dorrington et al., 2019; Toon et al., 2017; Williams et al., 2008).

A growing body of literature has focused on the diversification of Australian and New Guinea lineages with contemporary distributions on both sides of the Torres Strait. These systems, summarized in Joseph et al. (2019), have produced a variety of patterns from no diversification between Australian and New Guinea populations to multiple New Guinea lineages. *Coracina papuensis*, however, is the first system, to the best of our knowledge, demonstrating evidence for distinct lineages from both Oro and Central Province. Although, the individuals from both regions were recovered as monophyletic, we did not find strong support for the relationships of these clades relative to the other *C. papuensis* clades. Therefore, inferring multiple colonization events or *in situ* diversification in New Guinea remains difficult. Alternatively, an inability to recover well-resolved relationships may indicate a period of rapid diversification for *C. papuensis* that left too few synapomorphies to confidently reconstruct lineage relationships.

**Solomon Island Phylogeography**

The formation of the Solomon Islands is a result of a complex series of geologic events involving several phases of geologic activity beginning about 60 Mya and has continued into the present with volcanic activity (Neall & Trewick, 2008; Petterson et al., 1999). Fuchs, Cruaud, Couloux, and Pasquet (2007) estimated the Campephagidae to be less than 4 million years old, and estimated *C. papuensis* to have been independently evolving for only 500,000 years. Therefore, the complex geologic history of the Solomon Islands is likely to have had limited influence on the diversification of *C. papuensis*. Glacial cycling, on the other hand, appears to have had a significant influence, inhibiting avian diversification in the Solomon Islands. Bougainville, Choiseul, and Isabel were all connected during Glacial Maximum and together with Guadalcanal, separated by a narrow channel, have been called "Greater Bukida" (Mayr & Diamond, 2001). Although exceptions exist (Andersen et al., 2014; Andersen et al., 2013), many avian systems contain only slight or no differentiation among Bougainville, Choiseul, and Isabel. Furthermore, Guadalcanal populations are most often recovered either embedded within, or as the closest relative to a Bougainville, Choiseul and Isabel lineage (Andersen et al., 2015; Filardi & Smith, 2005; Smith & Filardi, 2007, Chapter 1). Current taxonomy for *C. papuensis* however, unites populations from Guadalcanal and the New Georgia Group as *C. p. elegens* and from Bougainville, Isabel and Choiseul as *C. p. perpadilla*. Nevertheless, we did not find support for this relationship. Instead, New Georgia Group individuals formed a clade, and the individual from Guadalcanal formed a well-supported clade with the individuals from Isabel and Choiseul (Fig 1). Thus, *C. papuensis* conforms to the well-established pattern of a single cohesive lineage inhabiting the Greater Bukida islands. Unfortunately, samples from Bougainville were unavailable, preventing us from fully sampling the range of this subspecies.

In total, we found support for three independent lineages within *C. papuensis* from the Solomon Islands. These include the two subspecies previously mentioned (i.e., *C. p. elegens* from the New Georgia group and *C. p. perpadilla* from Bougainville, Choiseul, Isabel and Guadalcanal) and a third taxon restricted to Malaita (i.e., *C. p. eyerdami*). In comparison, the *Rhipidura rufifrons* species complex, a well-known "great speciator" and understory species (Mayr & Diamond, 2001; Pratt, 2010) contains only two supported lineages across the same landscape (Chapter 1).

**Dispersal and Diversification:**

In theory and in many natural systems, gene flow prevents differentiation between populations. Therefore, lineages with higher dispersal ability should exhibit decreased diversification rates, a pattern identified by Smith et al. (2014) in continental avian systems. Continuous landscapes may offer fewer opportunities for isolation and therefore dispersal would likely further hinder diversification, whereas more fragmented landscapes such as island archipelagos, may depend on dispersal for colonization of novel regions and thus higher dispersal may promote diversification. Dispersal ability, however, is not a fixed trait and using foraging stratum may not be an appropriate assumption for some island taxa. For example, it is well known that for some plant systems, dispersal ability rapidly decreases after island colonization (Carlquist, 1966; Kavanagh & Burns, 2014). A similar pattern of rapid evolution of the traits underlying dispersal is highlighted by the bird genus *Zosterops* (Moyle et al., 2009) in which multiple species dispersed across hundreds of kilometers of ocean, but following their long distance dispersal did not remain homogenized by gene flow and instead differentiated, even for relatively proximate islands (2–20km). More explicitly, Pedersen et al. (2018) measured a morphological proxy for dispersal ability (hand-wing index) in *Edolisoma tenuirostre*, a close

relative of *C. papuensis*. They discovered that continental populations of *E. tenuirostre* had the highest dispersal ability, with archipelago and remote island populations both having a reduced ability to disperse. This hypothesis may help explain the pattern of diversification we observed for *C. papuensis*, whereby limited differentiation was observed across vast regions of continental habitat, but several independent lineages were observed within the comparatively small area of the Solomon Islands. Furthermore, not only did the three Solomon Island lineages all correspond to deep-water barriers, but *C. papuensis* is absent from Makira and the Santa Cruz islands, both of which were colonized by the understory species *Rhipidura rufifrons*. Therefore, island populations *C. papuensis* appear to be more dispersal limited than some understory species, perhaps indicating that inferring dispersal ability based on foraging stratum is not a valid assumption for *C. papuensis*.

# Chapter 3: Comparative phylogenomics of two closely related species complexes co-distributed throughout the Solomon Islands

## Introduction

The Monarch flycatchers (Monarchidae) are a clade of passerine birds found throughout the Old-World tropics. Despite their similarity to other "flycatchers" in terms of their ecomorphology (i.e., a dorso ventrally flattened bill and "sallying" feeding behavior) molecular data suggests the family Monarchidae belongs within the core Corvoidea (Barker, Cibois, Schikler, Feinstein, & Cracraft, 2004). Recently, work by Andersen et al. (2015) has helped to clarify relationships within the Monarchidae. The authors included 92% of the recognized species within the family, producing the most comprehensive phylogeny to date for the Monarch flycatchers. One of the four major clades identified by Andersen et al. (2015)—the "core-monarchs"—contained 10 genera and more than 65 taxa, including two radiations of Solomon Island monarchs: the chestnut-bellied monarchs and the black-and-white or pied monarchs. The taxa comprising these two radiations were taxa were considered congeners (Mayr & Diamond, 2001) until phylogenetic research by Filardi and Smith (2005) revealed non-monophyly and a deep genetic division within *Monarcha*. Based on this result, taxa within the genus were split between *Monarcha* and *Symposiachrus* (Christidis & Boles, 2008).

The chestnut-bellied radiation of monarchs from the Solomon Islands differentiated from their sister group approximately 0.4–0.98 million years ago (Uy, Cooper, & Chaves, 2019) and following the taxonomy of Clements et al. (2019) includes six taxa: *Monarcha castaneiventris* (4 subspecies), *Monarcha richardsii*, and *Monarcha erythrostictus*. *Monarcha richardsii* and *M. erythrostictus* are easily diagnosable: *M. richardsii* has a white crown and nape and *M. erythrostictus* has a white (male) or brown (female) crescent preceding the eye. *Monarcha richardsii* is endemic to the New Georgia Group and *M. erythrostictus* is commonly found throughout Bougainville, Buka and the Shortland Islands (Figure 6). The most phenotypically

diverse group within the chestnut-bellied monarchs are the four named taxa within *Monarcha castaneiventris*: *M. c. castaneiventris* (Choiseul, Isabel, Guadalcanal, and Malaita), *M. c. ugiensis* (Ugi, Santa Ana, Santa Catalina), M. c. *megarhynchus* (Makira), and *M. c. obscurior* (Russel Islands) (Figure 6). *Monarcha castaneiventris* includes multiple populations that do not possess the characteristic chestnut belly plumage and are instead all black. The melanic individuals on the Russel islands co-occur with chestnut-bellied individuals, but the populations on Santa Ana, Santa Catalina, Ugi, and Three Sister Islands are almost exclusively melanic. Genetic analyses did not, however, identify a "melanic clade" and instead indicated multiple evolutionary origins of melanism within *M. castaneiventris* (Uy et al., 2019; Uy et al., 2016).

The second radiation of Solomon Island monarchs—the black-and-white or pied monarchs—is comprised of a subset of the taxa originally contained within the "superspecies" *Monarcha (Symposiachrus) mandanensis*. Phylogenetic investigation rendered *M. mandanensis* a polyphyletic taxon (Filardi & Smith, 2005), but the Solomon Island taxa *Symposiachrus barbatus* (Bougainville, Isabel, Choiseul, Guadalcanal), *S. browni* (4 subspecies, New Georgia Group), *S. malaitae* (Malaita), and *S. vidua* (2 subspecies, Makira and Ugi) formed a clade (Andersen et al., 2015; Filardi & Smith, 2005). This group of black-and-white or pied monarchs varies in several plumage characteristics, including presence or absence of wing bars, scaling on the wings or chest, and a white cheek patch disjunct or connected to a white chest. Generally, other taxa outside of the Solomon Islands are included under the common name "pied monarchs" (Filardi & Smith, 2005), but for this manuscript pied monarchs will refer only to the populations from the Solomon Island archipelago (i.e., *S. barbatus*, *S. browni*, *S. malaitae*, and *S. vidua*).

The chestnut-bellied and pied monarchs present an opportunity to study two radiations that are closely related but radiated relatively recently, as neither complex is likely more than ~1

million years old (Filardi & Smith, 2005; Uy et al., 2019). The two radiations are broadly sympatric across the Solomon Island archipelago. The most substantial difference in distribution between the chestnut-bellied monarchs and pied is the absence of pied monarchs on the Russell Islands. Individuals from both complexes found on Guadalcanal, Choiseul and Isabel each constitute a single taxon (i.e., *M. c. castaneiventris* and *S. barbatus*); however, the complete range for each taxon are slightly different. Individuals from Malaita are included within *M. c. castaneiventris* and the distribution for S. barbatus also includes Bougainville. The New Georgia Group presents another contrast in distribution between the two radiations, with four taxa of pied monarchs described from these islands (i.e., *S. browni ganongae*, *S. b. nigrotectus*, *S. b. browni*, *S. b. meeki*) compared to only one chestnut-bellied monarch (i.e., *M. richardsii*).

Previous phylogeographic work has outlined major patterns in avian diversification across the Solomon Island archipelago (Andersen et al., 2014; Andersen et al., 2013; Pedersen et al., 2018; Smith & Filardi, 2007). For example, islands that have remained geographically isolated during periods of low sea level during Pleistocene glacial maxima, such as Malaita or Makira, frequently contain unique lineages. In contrast, islands that have recently experienced recent land bridge connections (e.g., Bougainville, Choiseul, and Isabel) often do not contain multiple lineages within a species complex. The chestnut-bellied and pied monarchs largely follow these major Solomon Island biogeographic patterns. Despite previous phylogenetic work, however, inter-lineage relationships within both complexes remain uncertain. Herein, we produce the first dataset of next-generation sequencing data for both the chestnut-bellied and pied monarchs. We evaluate genomic support for distinct lineages within both complexes, compare these findings to previous research, and place these complexes in context with other organisms that have diversified across the Solomon Islands.

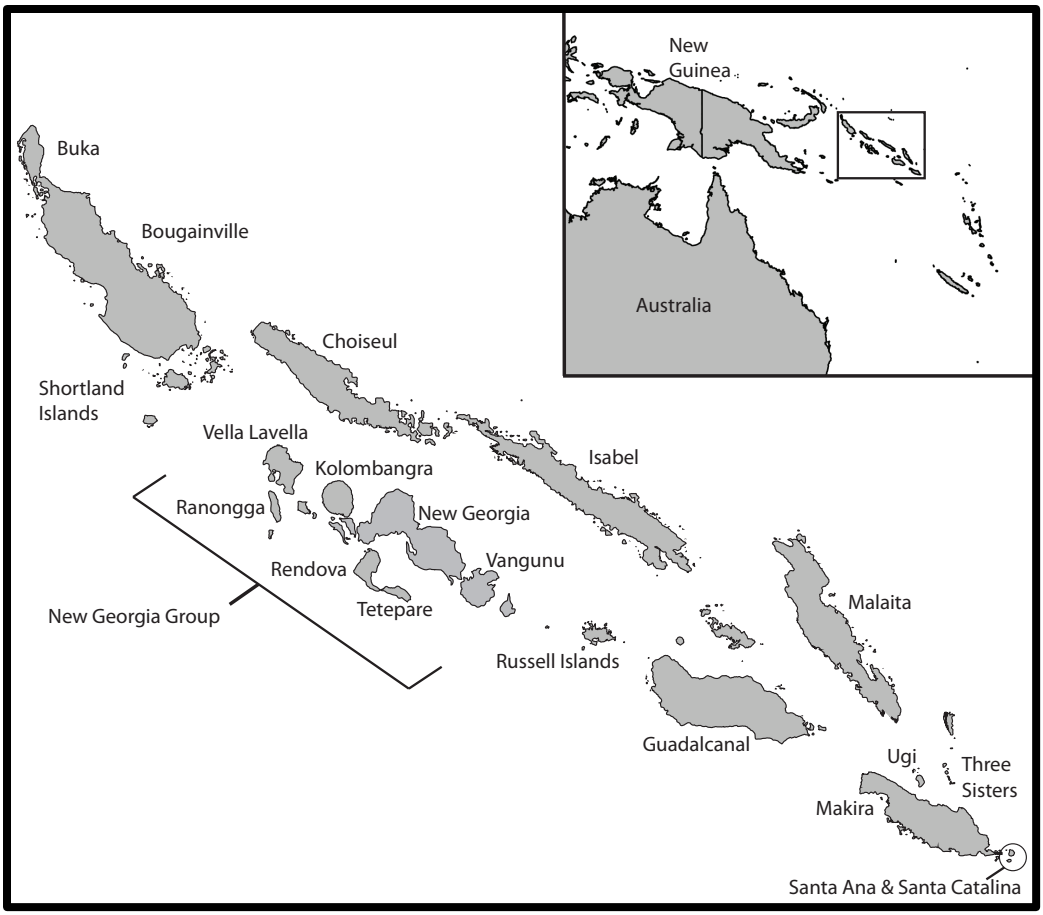


Figure 6. Map of the Solomon Islands

## Methods

Tissue samples were obtained for 24 individuals of chestnut-bellied monarchs and 21 individuals of pied monarchs (Table 3) from throughout the Solomon Islands. Total genomic DNA was extracted using Qiagen DNeasy tissue kits following the standard kit protocol. DNA concentration was determined with a Qubit 2.0 Fluorometer, and 500ng of DNA was sheared (fresh tissue samples only) using a Covaris S220 Sonicator with peak incident power of 175 W, 200 cycles per burst for 45 seconds, and 2% duty factor (Moyle et al., 2016). End repair, A-tailing, adapter ligation, and library amplification were completed using a Kapa Biosystems Library Prep kit. following the procedure of Faircloth et al. (2012) and using protocol modifications described by Moyle et al. (2016).

DNA libraries were merged into pools containing 8 equimolar samples for enrichment. Sequence capture and post-enrichment amplification followed standard protocols (Faircloth et al., 2012) using the Mycroarray MYbaits kit for Tetrapods UCE 5K version 1, which targets 5,060 ultra-conserved elements (UCE) loci. Libraries were sequenced for paired-end 100 bp sequencing on a partial lane of an Illumina HiSeq 2500 next-generation sequencer.

**Data Assembly**

De-multiplexing and removal of low-quality bases and adapter sequences of raw UCE reads was completed using llumniprocessor v2.0 (Del Fabbro, Scalabrin, Morgante, & Giorgi, 2013; Faircloth, 2013; Lohse et al., 2012). The Python package phyluce (Faircloth, 2016) enabled further data processing. Cleaned reads were assembled into contigs with the program Trinity (Grabherr et al., 2011) and contigs matching UCE loci were extracted for each taxon. The UCE loci were aligned with MAFFT (Katoh & Standley, 2013), allowing missing nucleotides on the margins of the alignment only if data were present for 65% of taxa included in the analysis. Alignments were then internally trimmed using Gblocks (Castresana, 2000; Talavera & Castresana, 2007) using default settings. Only those loci present in at least 50% of the samples were retained for analyses moving forward. To maximize the number of loci available for analysis, separate data matrices were produced for each species complex (i.e., for *Monarcha* and *Symposiachrus*).

Table 3. Locality and voucher information for *Monarcha* and *Symposiachrus* specimens included in Chapter 3

| Genus | Species | Institution | ID | UCE Loci | Locality |
|---|---|---|---|---|---|
| *Monarcha* | *castaneiventris castaneiventris* | AMNH | 6667 | 3607 | Isabel |
| *Monarcha* | *castaneiventris castaneiventris* | AMNH | 6686 | 3886 | Guadalcanal |

| Genus | Species | Institution | ID | UCE Loci | Locality |
|---|---|---|---|---|---|
| *Monarcha* | *castaneiventris castaneiventris* | AMNH | 21076 | 3713 | Guadalcanal |
| *Monarcha* | *castaneiventris castaneiventris* | UWBM | 69831 | 1983 | Isabel |
| *Monarcha* | *castaneiventris castaneiventris* | UWBM | 63306 | 2012 | Choiseul |
| *Monarcha* | *castaneiventris castaneiventris* | UWBM | 66030 | 2070 | Malaita |
| *Monarcha* | *castaneiventris castaneiventris* | UWBM | 66092 | 2073 | Malaita |
| *Monarcha* | *castaneiventris megarhynchus* | AMNH | 15303 | 3784 | Makira |
| *Monarcha* | *castaneiventris megarhynchus* | AMNH | 15344 | 3808 | Makira |
| *Monarcha* | *castaneiventris richardsii* | AMNH | 124362 | 2062 | Vella Lavella |
| *Monarcha* | *castaneiventris richardsii* | AMNH | 124257 | 2081 | Vella Lavella |
| *Monarcha* | *castaneiventris richardsii* | AMNH | 124307 | 2077 | Ranonga |
| *Monarcha* | *castaneiventris richardsii* | AMNH | 282 | 3771 | Kolombangara |
| *Monarcha* | *castaneiventris richardsii* | AMNH | 105 | 3814 | Gizo |
| *Monarcha* | *castaneiventris richardsii* | AMNH | 97 | 3748 | Gizo |
| *Monarcha* | *castaneiventris richardsii* | UWBM | 76391 | 2046 | Rendova |
| *Monarcha* | *castaneiventris richardsii* | UWBM | 76257 | 2081 | Rendova |
| *Monarcha* | *castaneiventris richardsii* | UWBM | 76383 | 2023 | Kolombangara |
| *Monarcha* | *castaneiventris richardsii* | UWBM | 66068 | 2031 | New Georgia |
| *Monarcha* | *castaneiventris richardsii* | UWBM | 66069 | 1998 | New Georgia |
| *Monarcha* | *castaneiventris richardsii* | UWBM | 76337 | 2082 | Tetepare |
| *Monarcha* | *castaneiventris ugiensis* | AMNH | 18933 | 3990 | Ugi |
| *Monarcha* | *castaneiventris ugiensis* | AMNH | 18934 | 3959 | Ugi |

| Genus | Species | Institution | ID | UCE Loci | Locality |
|---|---|---|---|---|---|
| *Monarcha* | *erythrostictus* | AMNH | 226139 | 2042 | Buka |
| *Symposiachrus* | *barbatus barbatus* | AMNH | 21075 | 3882 | Guadalcanal |
| *Symposiachrus* | *barbatus barbatus* | AMNH | 6647 | 2990 | Isabel |
| *Symposiachrus* | *barbatus barbatus* | AMNH | 6669 | 3805 | Isabel |
| *Symposiachrus* | *barbatus barbatus* | UWBM | 60245 | 2006 | Guadalcanal |
| *Symposiachrus* | *barbatus barbatus* | UWBM | 63163 | 2027 | Choiseul |
| *Symposiachrus* | *barbatus barbatus* | UWBM | 63161 | 1999 | Choiseul |
| *Symposiachrus* | *barbatus malaitae* | AMNH | 21020 | 3348 | Malaita |
| *Symposiachrus* | *barbatus malaitae* | AMNH | 21042 | 3731 | Malaita |
| *Symposiachrus* | *browni browni* | AMNH | 225529 | 1984 | Vangunu |
| *Symposiachrus* | *browni browni* | AMNH | DOT18905 | 2080 | New Georgia |
| *Symposiachrus* | *browni browni* | AMNH | 188 | 3916 | Kolombangara |
| *Symposiachrus* | *browni browni* | AMNH | 189 | 3897 | Kolombangara |
| *Symposiachrus* | *browni browni* | UWBM | 76420 | 2057 | New Georgia |
| *Symposiachrus* | *browni ganongae* | AMNH | 18952 | 3872 | Ranongga |
| *Symposiachrus* | *browni meeki* | AMNH | 18908 | 3666 | Rendova |
| *Symposiachrus* | *browni meeki* | UWBM | 76264 | 2041 | Rendova |
| *Symposiachrus* | *browni nigrotectus* | AMNH | 115 | 3780 | Vella Lavella |
| *Symposiachrus* | *browni nigrotectus* | AMNH | 123 | 4109 | Vella Lavella |
| *Symposiachrus* | *vidua squamulatus* | AMNH | 17172 | 3018 | Ugi |
| *Symposiachrus* | *vidua vidua* | AMNH | 15323 | 4020 | Makira |
| *Symposiachrus* | *vidua vidua* | AMNH | 15335 | 3845 | Makira |

AMNH American Museum of Natural History, UWBM University of Washington Burke Museum

**Phylogenetic Analysis**

The two 50% complete datasets of concatenated UCE loci (i.e., one for each complex) were analyzed using RAxML v8 (Stamatakis, 2014) to infer phylogenetic relationships among individuals in a maximum likelihood (ML) framework. For both datasets, sequence evolution was modeled using the GTR + G model. In most phylogenetic scenarios, the most common gene tree topology will match the underlying species tree. In some situations (e.g., with two or more short successive internodes), however, the most common gene tree will be in conflict with the true species tree, a situation referred to as the "anomaly zone" (Degnan & Rosenberg, 2006; Degnan & Salter, 2005). Under this scenario, concatenation of loci will support an inaccurate phylogeny. To avoid complications caused by the anomaly zone, species tree topologies were inferred using SVDquartets (Chifman & Kubatko, 2014) and ASTRAL (Mirarab et al., 2014), both of which are gene-tree based coalescent methods. SVDquartets was completed in PAUP* (Swofford, 1998), sampling all possible quartets and establishing nodal support with 500 bootstrap replicates. Quartet trees were integrated into a single species tree with PAUP* using the *Quartet FM* algorithm (Reaz et al., 2014). For ASTRAL, individual gene trees were generated with RAxML v 8.1.3 and nodal support was established from quartet frequencies (Sayyari & Mirarab, 2016). Based on previous work, the phylogenetic analyses were rooted with *M. cinerascens* for the chestnut-bellied monarchs and with *S. trivigatus* and *S. guttula* for the pied monarchs (Andersen et al., 2015).

## Results

On average, 2,181 and 4,024 UCE loci were enriched per sample for the individuals from the chestnut-bellied and pied monarch complexes, respectively. The 50% complete dataset for the chestnut-bellied monarchs had an average locus length of 526.73 base pairs (bp) (175–1,334

bp) and contained 2,978 informative characters. The 50% complete dataset for the pied monarchs had an average locus length of 564.72 bp (206–1,528 bp) and contained 6,439 informative characters. Datasets for the two complexes were analyzed separately rather than reduce each dataset to only the ~2,000 UCE loci shared between both.

**Chestnut-bellied Monarchs: RAxML and SVDquartets Results**

The topologies resulting from the RAxML and SVDquartets analyses for species of *Monarcha* were mostly congruent with one another, with no strongly-supported relationships in conflict between analyses (Figure 7; RAxML tree shown). The individuals from Ugi (i.e., *M. c. ugiensis*) and Makira (i.e., *M. c. megarhynchus*) formed a well-supported clade, with individuals from the two islands forming reciprocally monophyletic groups (Figure 7, clade M1). These Ugi and Makira populations were sister to a clade containing the individuals from Choiseul, Isabel, Guadalcanal and Malaita (i.e., *M. c. castaneiventris*, Figure 7, clade M2), albeit with relatively low support in both analyses (BS = 70 and BS = 51 for RAxML and SVDquartets, respectively). Within the M2 clade, the two individuals from Malaita (i.e., *M. c. malaitae*) were recovered as sister to one another with strong support (BS = 100 in both RAxML and SVDquartets analyses) but their placement within M2 clade was unresolved. Both analyses found strong support for a clade uniting all individuals from the New Georgia Group (i.e., *M. richardsii*) with the sample from Buka Island (i.e., *M. erythrostictus*), but few relationships within the clade where well-supported (Figure 7, M3 clade).

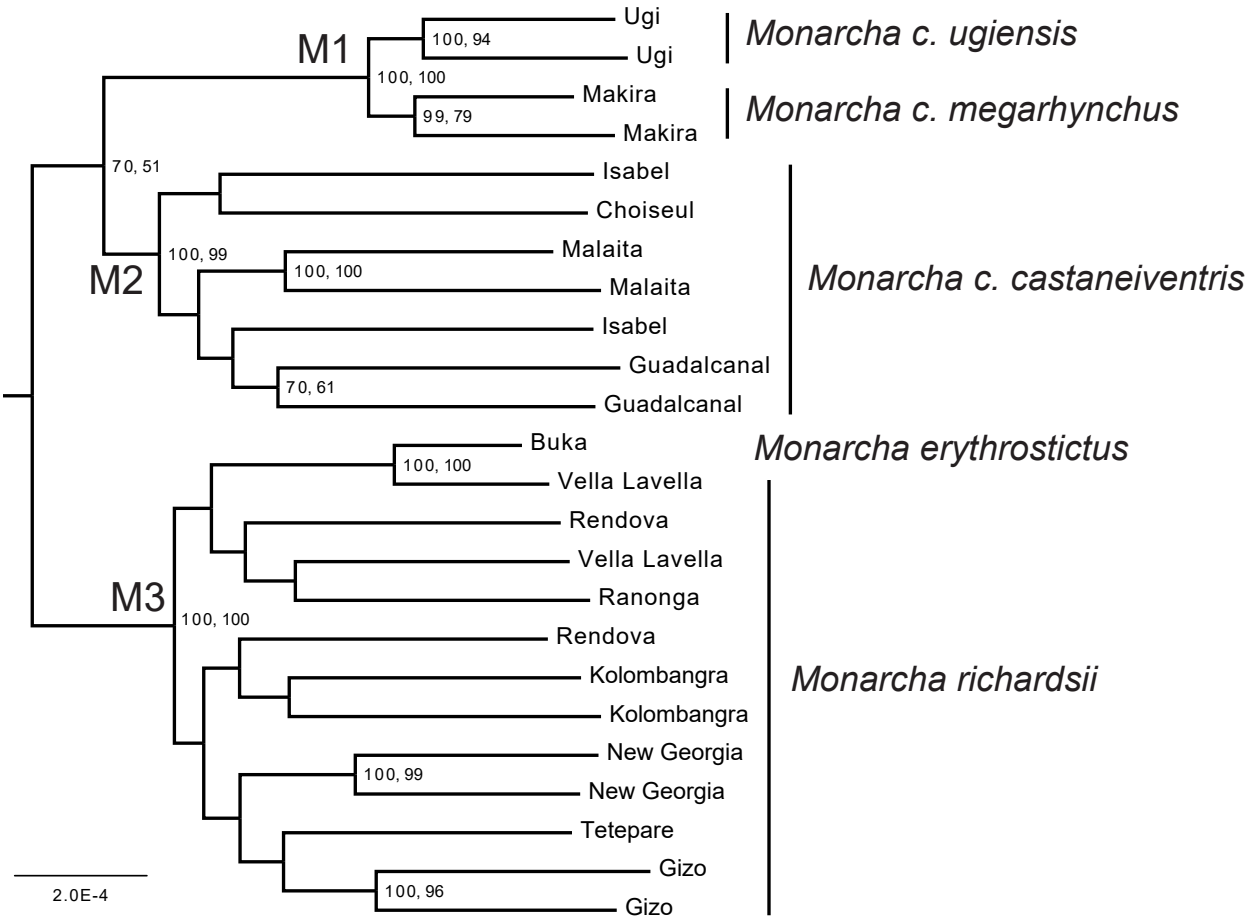


Figure 7. The maximum likelihood analysis with RAxML on the concatenated 2181 UCE loci. Tips are labeled by geographic origin and current taxonomy shown on the left. The three major clades M1–M3 are labeled at the common ancestor to the clade. Node support was established by bootstrapping and the first number corresponds to RAxML analysis and the second shows the support from the species tree analysis SVDquartets. Only nodes receiving more than 50 bs support for both analyses are shown

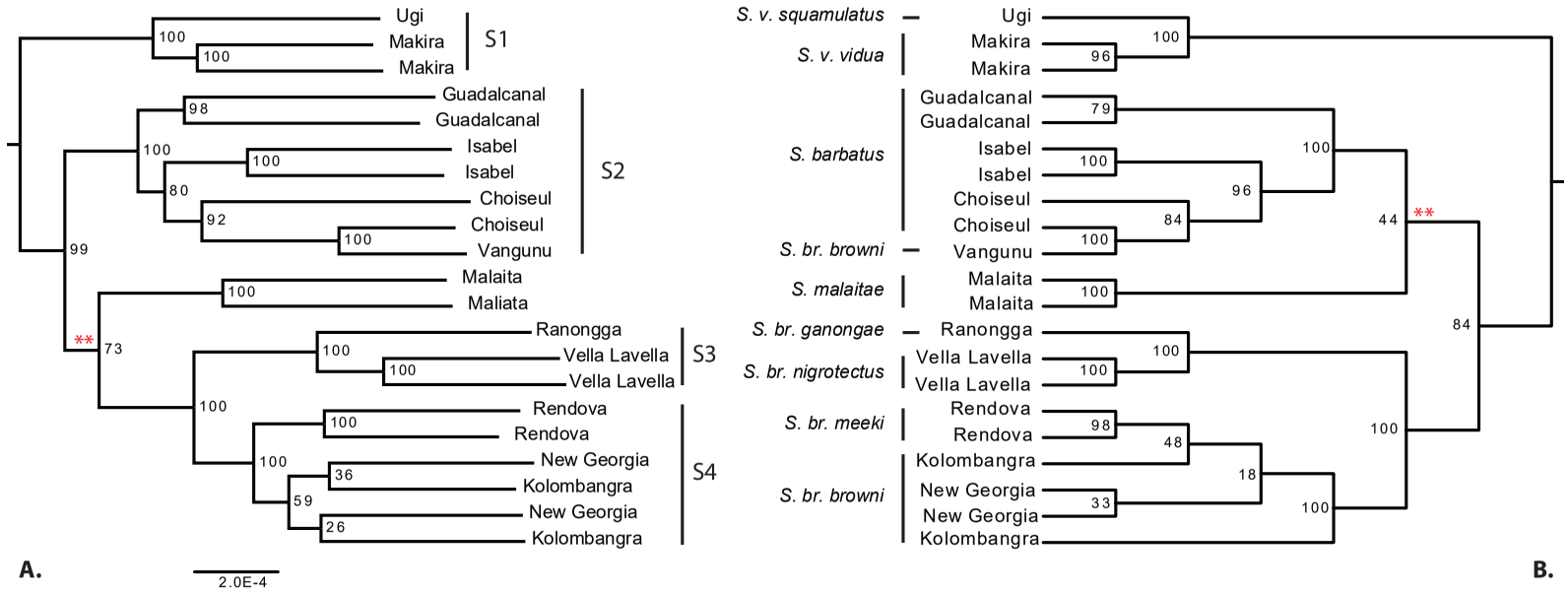


Figure 8. The four major clades S1–S4 are labeled on tree A and the current taxonomy is indicated on tree B. The red asterisks indicate the incongruent relationships recovered by the RAxML and SVDquartets analyses for the individuals from Malaita. A) Maximum likelihood phylogeny estimated with RAxML. Nodal support was determined by completing 500 rapid bootstraps. B) Species tree phylogeny estimated with SVDquartets considering all possible quartet trees. Nodal support was determined by completing 500 bootstrap replicates.

**Pied Monarchs: RAxML and SVDquartets Results**

Phylogenetic analysis of the *Symposiachrus* dataset using RAxML and SVDquartets produced few incongruent relationships. In both analyses, individuals from Makira (i.e., *S. v. vidua*) and Ugi (i.e., *S. v. squamulatus*) formed a clade sister to the remaining individuals from the Solomon Island (BS = 99 and BS = 84 for RAxML and SVDquartets, respectively) (Figure 8, S1 clade). Both methods also recovered strong support for a clade containing individuals from Guadalcanal, Isabel, and Choiseul (i.e., *S. barbatus*, Figure 8, S2 clade). Within the S2 clade, individuals from Guadalcanal and Isabel were well-supported as independent lineages, sequentially sister to the clade containing individuals from Choiseul and Vangunu (i.e., *S. browni*).

With the exception of the sample from Vangunu, individuals from the New Georgia Group (i.e., *S. browni*) formed two well-supported, reciprocally monophyletic groups (Figure 8, S3 and S4 clades). The S3 clade contained individuals from northern New Georgia Group islands (i.e., *S. b. ganongae* from Ranongga and *S. b. nigrotectus* from Vella Lavella) and the S4 clade contained individuals from the southern New Georgia Group islands (i.e., *S. b. meeki* from Rendova, and *S. b. browni* from New Georgia and Kolombangra). The sole incongruence between analyses produced using the two methods involved the placement of the two individuals from Malaita (i.e., *S. b. malaitae*): RAxML analyses supported a sister relationship between individuals of *S. b. malaitae* from Malaita and individuals of *S. browni* from the New Georgia Group in the S3 and S4 clades (BS = 73), whereas SVDquartets placed the two individuals of *S. b. malaitae* sister to clade S2 (BS = 44).

**Chestnut-Bellied and Pied Monarchs: ASTRAL Results**

In contrast to RAxML and SVDquartets, ASTRAL requires as input individual gene tree topologies that are assumed to be accurately estimated. Datasets for the chestnut-bellied and pied monarchs had an average of 1.6 and 1.37 parsimony informative characters per locus, respectively; thus, the assumption of well-resolved gene trees was problematic. Given such low numbers of informative sites per locus, it was unsurprising that relationships within the chestnut-bellied or pied monarchs could not be accurately estimated. As ASTRAL assumes gene trees estimated without error, the ASTRAL analyses presumably produced a phylogeny with generally high nodal support; however, the relationships inferred by ASTRAL were largely incongruent with those inferred by RAxML and SVDquartets. Given that ASTRAL was able to produce well-resolved gene trees for either dataset, these results will not be discussed.

## Discussion

With several thousand loci sampled from throughout the genome, the data collected and analyzed in this study increased the resolution of interrelationships within both the chestnut-bellied and pied monarch complexes and revealed new evidence of admixture among monarch taxa. Similar to previous investigations that included individuals of *Monarcha* and *Symposiachrus*, several clades received strong phylogenetic support. The relationships inferred among clades were occasionally not well-supported (e.g., clades M1 and M2), but overall were more well-resolved herein as compared to previous investigations. The different phylogenetic analyses used in this study produced incongruent results for the relationship of pied monarchs from Malaita relative to the rest of the complex, and the hypothesized relationship recovered for chestnut-bellied monarchs from Malaita relative to the other chestnut-bellied monarchs did not align with the relationships previously hypothesized for this species complex.

### Chestnut-Bellied Monarchs

Recent phylogenetic studies of *M. castaneiventris* using between two and five loci (Andersen et al., 2015; Uy et al., 2019) have generally produced well-supported relationships within the group. Here, using 2,248 loci, we failed to recover well-supported relationships between the three major clades (i.e., M1–M3). After accounting for differences in sampling effort, previous work and our study converge on a similar number of independent lineages within the chestnut-bellied monarch species complex, but significant differences emerge for the inferred relationships within those lineages, revealing a pattern of mito-nuclear discordance. For example, our results based on a UCE dataset found individuals from Malaita embedded within a clade containing individuals from Isabel, Choiseul and Guadalcanal, while Uy et al. (2019) found strong support for individuals from Malaita as sister to all the other individuals using predominantly mitochondrial loci. Similarly, (Andersen et al., 2015) analyzed 5 independent loci (four nuclear and one mitochondrial; ~1,200 parsimony informative characters) and was unable to resolve the position of the Malaitan population within the chestnut-bellied species complex more broadly. Incongruence between results based on mitochondrial versus nuclear datasets is not uncommon, and can arise for several reasons, including sex-biased hybridization, sex-biased dispersal, incomplete lineage sorting, and selective sweeps within mitochondrial DNA (Campillo et al., 2018; Toews & Brelsford, 2012).

For many avian systems in the Solomon Islands, the populations on Bougainville are found either sister to, or embedded within, a lineage containing individuals from Choiseul and Isabel. These three islands have been repeatedly connected to one another by land bridges during periods of low sea level, likely facilitating gene flow, and thus limiting differentiation, between their respective populations. Uy et al. (2019), however, found that the chestnut-bellied complex diverges from this trend. A subset of the individuals from Choiseul formed a clade with

individuals from Isabel and Guadalcanal, but a pair of samples formed a clade with individuals from Bougainville that was collectively sister to a clade of individuals from the New Georgia Group, albeit with low support (posterior probability = 57). The majority of the phylogenetic information used by Uy et al. (2019) came from the haploid and maternally-inherited mitochondrial genome. Therefore, low levels of gene flow from Bougainville to Choiseul could introduce mitochondrial haplotypes predominantly found in Bougainville, explaining their observation on Choiseul. Here, we recovered an individual from Buka, an island separated from Bougainville by a channel about ¼ km across, embedded with strong support within the clade of individuals from the New Georgia Group. Unfortunately, with only one sample from Buka and one sample from Choiseul, we are unable to evaluate levels of gene flow between populations or test the monophyly of populations from Choiseul.

**Pied Monarchs**

Previous research by (Andersen et al., 2015) identified several well-supported and geographically cohesive lineages within the pied monarchs. The relationships among these lineages, however, were not well-supported. Analysis of genome-wide UCE loci in this study resolved some of these relationships, but others remain unclear. For example, although previous datasets could not confidently place individuals from Makira and Ugi within the pied monarch complex, analysis of UCE loci unambiguously placed these individuals (S1 clade) as the sister group to all other pied monarch populations. This result, however, was not well-supported in our analyses, or in those of previous investigations (Andersen et al., 2015; Filardi & Smith, 2005).

Unlike in the chestnut-bellied monarch (i.e., *M. castaneiventris*), rufous fantail (i.e., *Rhipidura rufifrons*), or white-bellied cuckooshrike (i.e., *Coracina papuensis*) species complexes, the pied monarch species complex contains multiple taxa described from the New

Georgia Group. Few avian systems have diversified into multiple lineages across this landscape, with other notable exceptions within the *Zosterops* (Moyle et al. in review.), *Pachycephala* (Andersen et al., 2014), and *Myzomela* (Clements et al., 2019) complexes. The pied monarchs contain four described subspecies: *S. b. browni* (New Georgia and Kolombangra), *S. b. meeki* (Rendova and Tetepare), *S. b. nigrotectus* (Vella Lavella), and *S. b. ganongae* (Ranongga). Andersen et al. (2015) provided the first molecular evidence supporting multiple lineages of *Symposiachrus* within the New Georgia Group, but the relationships among these lineages were unresolved, making differentiation between *in situ* diversification and multiple colonization events of the New Georgia Group impossible. Here, we found overall support for a single common ancestor for all lineages of *Symposiachrus* from the New Georgia Group (Figure 8); however, we cannot eliminate the possibility of more complex scenarios involving multiple dispersal events coupled with extinction of the founding population. In fact, the genomic data in the present study indicate a diversification history perhaps more complex than a single colonization event followed by allopatric differentiation.

We identified the individual from Vangunu (in the New Georgia Group) to be unambiguously embedded within the S2 clade (Figure 8), a pattern suggestive of recent gene flow between populations from the New Georgia Group and other islands. As this result is based only on a single sample, however, we caution against its use to infer the demographic history for pied monarchs in the Solomon Islands. The DNA extracted from this single individual from Vangunu was isolated from a museum study skin (i.e., from a toepad), and since recent investigations have indicated that historical samples are more prone to contamination than fresh tissue samples (Moyle, Hosner, Jones, & Outlaw, 2015), the quality of the data produced for this sample are further in question. Finally, UCE loci are best suited to estimate relationships among

distinct evolutionary lineages, rather than to estimate population genetic parameters. Therefore, to evaluate the population demography of pied monarchs, future research should focus on generating full genomic, rather than reduced-representation, datasets.

## Conclusions

Despite sharing similar life histories, lineage ages, and geographic distributions, the chestnut-bellied and pied monarch species complexes share few phylogeographic patterns. Overall, we did not identify strongly-supported structure for *M. c. castaneiventris* populations distributed across Guadalcanal, Choiseul, Isabel. In contrast—setting aside the single, questionable individual from Vangunu—populations of *S. barbatus* from each island on which it is found (i.e., Guadalcanal, Choiseul, Isabel) formed reciprocally monophyletic groups. Future studies on pied monarchs would benefit from increased sampling of individuals from these islands, but also from the inclusion of representatives from Bougainville and the Florida Islands.

One commonality between the two species complexes studied herein is that both the chestnut-bellied and pied monarchs have diversified into two lineages across Makira and Ugi (Figure 7, M1 clade and Figure 8, S1 clade, respectively). The broader phylogeographic patterns exhibited by the two clades, however, deviate from one another. Clade S1 is sister to the rest of the taxa from the Solomon Islands, but the closest relative to clade M1 is *M. c. castaneiventris*. Based on the results presented herein, further investigation into the population structure for individuals from the New Georgia Group is warranted for both systems. The chestnut-bellied species complex previously showed no evidence for multiple taxa from the New Georgia Group, a result further supported by this study, but we did identify a single individual from Bougainville (i.e., *M. erythrostictus*) as a phenotypically-distinct population embedded within the New Georgia Group clade (*M. richardsii*). For the pied monarchs, we found a single individual from

the New Georgia Group that was more genetically similar to individuals from Choiseul than to other individuals from the New Georgia Group, indicating recent gene flow between populations from the New Georgia Group and Choiseul. Future research on the pied monarchs from the Solomon Islands should focus on generating datasets more suitable than UCEs for determining demographic parameters like gene flow, such as restriction enzyme-associated DNA sequencing (RAD-seq) or genome resequencing.

## References


Allison, A., & Leisz, S. (2009). Analysis of the impacts of climate change on the herpetofauna of the Papuan region (New Guinea to the Solomon Islands). Bishop Museum, Honolulu.

Andersen, M. J., Hosner, P. A., Filardi, C. E., & Moyle, R. G. (2015). Phylogeny of the monarch flycatchers reveals extensive paraphyly and novel relationships within a major Australo-Pacific radiation. *Mol. Phylogenet. Evol., 83*(C), 118-136.

Andersen, M. J., Nyári, Á. S., Mason, I., Joseph, L., Dumbacher, J. P., Filardi, C. E., & Moyle, R. G. (2014). Molecular systematics of the world's most polytypic bird: the *Pachycephala pectoralis/ melanura* (Aves: Pachycephalidae) species complex. *Zoological Journal of the Linnean Society, 170*(3), 566–588.

Andersen, M. J., Oliveros, C. H., Filardi, C. E., & Moyle, R. G. (2013). Phylogeography of the Variable Dwarf-Kingfisher *Ceyx Lepidus* (Aves: Alcedinidae) inferred from mitochondrial and nuclear DNA sequences. *Auk, 130*(1), 118–131.

Andersen, M. J., Shult, H. T., Cibois, A., Thibault, J. C., Filardi, C. E., & Moyle, R. G. (2015). Rapid diversification and secondary sympatry in Australo-Pacific kingfishers (Aves: Alcedinidae: *Todiramphus*). *Royal Society Open Science, 2*(2), 140375–140375.

Andolfatto, P., Davison, D., Erezyilmaz, D., Hu, T. T., Mast, J., Sunayama-Morita, T., & Stern, D. L. (2011). Multiplexed shotgun genotyping for rapid and efficient genetic mapping. *Genome Research, 21*(4), 610–617.

Barker, F. K., Cibois, A., Schikler, P., Feinstein, J., & Cracraft, J. (2004). Phylogeny and diversification of the largest avian radiation. *Proceedings of the National Academy of Sciences of the United States of America, 101*(30), 11040-11045.

Belliure, J., Sorci, G., Møller, A., & Clobert, J. (2000). Dispersal distances predict subspecies richness in birds. *Journal of Evolutionary Biology, 13*(3), 480-487.

Boles, W., & Christi, D. (2019). Rufous Fantail (*Rhipidura rufifrons*). In: del Hoyo, J., Elliott, A., Sargatal, J., Christie, D.A. & de Juana, E. (eds.). Handbook of the Birds of the World Alive. Lynx Edicions, Barcelona.

Bolmgren, K., & Eriksson, O. (2005). Fleshy fruits–origins, niche shifts, and diversification. *Oikos, 109*(2), 255-272.

Campagna, L., Gronau, I., Silveira, L. F., Siepel, A., & Lovette, I. J. (2015). Distinguishing noise from signal in patterns of genomic divergence in a highly polymorphic avian radiation. *Molecular Ecology, 24*(16), 4238–4251.

Campillo, L., Manthey, J. D., Thomson, R. C., Hosner, P. A., & Moyle, R. G. (2019). Genomic differentiation in an endemic Philippine genus (Aves: Sarcophanops) due to geographic isolation on recently disassociated islands. *PeerJ Preprints, 7*, e27546v27541.

Campillo, L. C., Oliveros, C. H., Sheldon, F. H., & Moyle, R. G. (2018). Genomic data resolve gene tree discordance in spiderhunters (Nectariniidae, Arachnothera). *Molecular Phylogenetics and Evolution, 120*, 151-157.

Carlquist, S. (1966). The biota of long-distance dispersal. II. Loss of dispersibility in Pacific Compositae. *Evolution, 20*(1), 30-48.

Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution, 17*(4), 540-552.

Catchen, J. M., Amores, A., Hohenlohe, P., Cresko, W., & Postlethwait, J. H. (2011). Stacks: building and genotyping loci de novo from short-read sequences. *G3: Genes, Genomes, Genetics, 1*(3), 171–182.

Chifman, J., & Kubatko, L. (2014). Quartet inference from SNP data under the coalescent model. *Bioinformatics, 30*(23), 3317-3324.

Christidis, L., & Boles, W. (2008). *Systematics and taxonomy of Australian birds*: Csiro Publishing.

Cibois, A., Beadell, J. S., Graves, G. R., Pasquet, E., Slikas, B., Sonsthagen, S. A., . . . Fleischer, R. C. (2011). Charting the course of reed-warblers across the Pacific islands. *Journal of Biogeography, 38*(10), 1963–1975.

Cibois, A., Thibault, J. C., & Pasquet, E. (2007). Uniform phenotype conceals double colonization by reed-warblers of a remote Pacific archipelago. *Journal of Biogeography, 34*(7), 1150–1166.

Claramunt, S., Derryberry, E. P., Remsen, J., & Brumfield, R. T. (2012). High dispersal ability inhibits speciation in a continental radiation of passerine birds. *Proceedings of the Royal Society of London B: Biological Sciences, 279*(1733), 1567-1574.

Clements, J. F., Schulenberg, T. S., Iliff, M. J., Billerman, S. M., Fredericks, T. A., Sullivan, B. L., & Wood, C. L. (2019). The eBird/Clements Checklist of Birds of the World: v2019. Retrieved from http://www.birds.cornell.edu/clementschecklist/download/

Darwin, C. (1859). *On the origin of species*. London: John Murray.

Degnan, J. H., & Rosenberg, N. A. (2006). Discordance of species trees with their most likely gene trees. *PLoS Genetics, 2*(5), e68.

Degnan, J. H., & Salter, L. A. (2005). Gene tree distributions under the coalescent process. *Evolution, 59*(1), 24-37.

Del Fabbro, C., Scalabrin, S., Morgante, M., & Giorgi, F. M. (2013). An extensive evaluation of read trimming effects on Illumina NGS data analysis. *PloS one, 8*(12), e85024.

Diamond, J. M. (1974). Colonization of exploded volcanic islands by birds: the supertramp strategy. *Science, 184*(4138), 803–806.

Diamond, J. M., Gilpin, M. E., & Mayr, E. (1976). Species-distance relation for birds of the Solomon Archipelago, and the paradox of the great speciators. *Proceedings of the National Academy of Sciences, 73*(6), 2160–2164.

Dorrington, A., Joseph, L., Hallgren, W., Mason, I., Drew, A., Hughes, J. M., & Schmidt, D. J. (2019). Phylogeography of the blue-winged kookaburra Dacelo leachii across tropical northern Australia and New Guinea. *Emu-Austral Ornithology*, 1-13.

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology, 14*(8), 2611–2620.

Fabre, P. H., Moltensen, M., Fjeldså, J., Irestedt, M., Lessard, J. P., & Jønsson, K. A. (2013). Multiple waves of colonization by monarch flycatchers (*Myiagra*, *Monarchidae*) across the Indo-Pacific and their implications for coexistence and speciation. *Journal of Biogeography, 41*(2), 274-286.

Faircloth, B. C. (2013). Illumiprocessor: a trimmomatic wrapper for parallel adapter and quality trimming. *See http://dx. doi. org/10.6079/J9ILL (accessed 4 November 2016)*.

Faircloth, B. C. (2016). PHYLUCE is a software package for the analysis of conserved genomic loci. *Bioinformatics, 32*, 786-788. doi:10.1093/bioinformatics/btv646

Faircloth, B. C., McCormack, J. E., Crawford, N. G., Harvey, M. G., Brumfield, R. T., & Glenn, T. C. (2012). Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. *Systematic Biology, 61*(5), 717-726.

Filardi, C. E., & Moyle, R. G. (2005). Single origin of a pan-Pacific bird group and upstream colonization of Australasia. *Nature, 438*(7065), 216-219.

Filardi, C. E., & Smith, C. E. (2005). Molecular phylogenetics of monarch flycatchers (genus *Monarcha*) with emphasis on Solomon Island endemics. *Molecular Phylogenetics and Evolution, 37*(3), 776-788.

Fuchs, J., Cruaud, C., Couloux, A., & Pasquet, E. (2007). Complex biogeographic history of the cuckoo-shrikes and allies (Passeriformes: Campephagidae) revealed by mitochondrial and nuclear sequence data. *Mol. Phylogenet. Evol., 44*(1), 138-153.

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., . . . Zeng, Q. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology, 29*(7), 644-652.

Harvey, M. G., Seeholzer, G. F., Smith, B. T., Rabosky, D. L., Cuervo, A. M., & Brumfield, R. T. (2017). Positive association between population genetic differentiation and speciation rates in New World birds. *Proceedings of the National Academy of Sciences, 114*(24), 6328-6333.

Irestedt, M., Fabre, P. H., Batalha-Filho, H., Jønsson, K. A., Roselaar, C. S., Sangster, G., & Ericson, P. G. (2013). The spatio-temporal colonization and diversification across the Indo-Pacific by a 'great speciator'(Aves, *Erythropitta erythrogaster*). *Proceedings of the Royal Society of London B: Biological Sciences, 280*(1759), 20130309.

Jombart, T. (2008). adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics, 24*(11), 1403–1405.

Jombart, T., & Ahmed, I. (2011). adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics, 27*(21), 3070–3071.

Jombart, T., Devillard, S., & Balloux, F. (2010). Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics, 11*(1), 94.

Jønsson, K. A., Bowie, R. C., Nylander, J. A., Christidis, L., Norman, J. A., & Fjeldså, J. (2010). Biogeographical history of cuckoo-shrikes (Aves: Passeriformes): transoceanic colonization of Africa from Australo-Papua. *Journal of Biogeography, 37*(9), 1767-1781.

Jønsson, K. A., Irestedt, M., Christidis, L., Clegg, S. M., Holt, B. G., & Fjeldså, J. (2014). Evidence of taxon cycles in an Indo-Pacific passerine bird radiation (Aves: *Pachycephala*). *Proceedings of the Royal Society of London B: Biological Sciences, 281*(1777), 20131727.

Joseph, L., Bishop, K. D., Wilson, C. A., Edwards, S. V., Iova, B., Campbell, C. D., . . . Drew, A. (2019). A review of evolutionary research on birds of the New Guinean savannas and closely associated habitats of riparian rainforests, mangroves and grasslands. *Emu, 119*(3), 317-330. doi:https://doi.org/10.1080/01584197.2019.1615844

Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution, 30*(4), 772-780.

Kavanagh, P. H., & Burns, K. C. (2014). The repeated evolution of large seeds on islands. *Proceedings of the Royal Society of London B: Biological Sciences, 281*(1786), 20140675.

Kearns, A. M., Joseph, L., & Cook, L. G. (2013). A multilocus coalescent analysis of the speciational history of the Australo-Papuan butcherbirds and their allies. *Molecular Phylogenetics and Evolution, 66*(3), 941–952.

Kennedy, J. D., Borregaard, M. K., Jønsson, K. A., Marki, P. Z., Fjeldså, J., & Rahbek, C. (2016). The influence of wing morphology upon the dispersal, geographical distributions and diversification of the Corvides (Aves; Passeriformes). *Proceedings of the Royal Society B: Biological Sciences, 283*(1844), 20161922.

Koblmüller, S., Egger, B., Sturmbauer, C., & Sefc, K. M. (2010). Rapid radiation, ancient incomplete lineage sorting and ancient hybridization in the endemic Lake Tanganyika cichlid tribe *Tropheini*. *Molecular Phylogenetics and Evolution, 55*(1), 318–334.

Linck, E., Schaack, S., & Dumbacher, J. P. (2016). Genetic differentiation within a widespread "supertramp" taxon: Molecular phylogenetics of the Louisiade White-eye (*Zosterops griseotinctus*). *Molecular Phylogenetics and Evolution, 94*, 113–121.

Lohse, M., Bolger, A. M., Nagel, A., Fernie, A. R., Lunn, J. E., Stitt, M., & Usadel, B. (2012). R obi NA: A user-friendly, integrated software solution for RNA-Seq-based transcriptomics. *Nucleic Acids Research, 40*(W1), W622-W627.

Losos, J. B., Jackman, T. R., Larson, A., de Queiroz, K., & Rodríguez-Schettino, L. (1998). Contingency and determinism in replicated adaptive radiations of island lizards. *Science, 279*(5359), 2115–2118.

MacArthur, R. H., & Wilson, E. O. (1963). An equilibrium theory of insular zoogeography. *Evolution, 17*(4), 373–387.

MacArthur, R. H., & Wilson, E. O. (1967). *The Theory of Island Biogeography*. Princeton: Princeton Univ. Press.

Manthey, J. D., Campillo, L. C., Burns, K. J., & Moyle, R. G. (2016). Comparison of Target-capture and Restriction-site Associated DNA Sequencing for Phylogenomics: a Test in Cardinalid Tanagers (Aves, Genus: *Piranga*). *Systematic Biology, 65*(4), 640–650.

Manthey, J. D., Moyle, R. G., Gawin, D. F., Rahman, M. A., Ramji, M. F. S., & Sheldon, F. H. (2017). Genomic phylogeography of the endemic Mountain Black-eye of Borneo (Chlorocharis emiliae): montane and lowland populations differ in patterns of Pleistocene diversification. *Journal of Biogeography, 44*(10), 2272-2283.

Mayr, E. (1941). The origin and the history of the bird fauna of Polynesia. *Proc. Sixth Pacific Sci. Congr. Zool, 4*, 197-216.

Mayr, E. (1942). *Systematics and The Origin of Species, From The Viewpoint of A Zoologist*: Harvard University Press.

Mayr, E., & Diamond, J. (2001). *The Birds of Northern Melanesia: Speciation, Ecology, and Biogeography*. New York: Oxford Univ. Press.

Miller, M. R., Dunham, J. P., Amores, A., Cresko, W. A., & Johnson, E. A. (2007). Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. *Genome Research, 17*(2), 240–248.

ASTRAL: genome-scale coalescent-based species tree estimation., 17, 30 Cong. Rec. i541-548 (2014).

Moyle, R. G., Filardi, C. E., Smith, C. E., & Diamond, J. (2009). Explosive Pleistocene diversification and hemispheric expansion of a "great speciator". *Proceedings of the National Academy of Sciences, 106*(6), 1863–1868.

Moyle, R. G., Hosner, P. A., Jones, A. W., & Outlaw, D. C. (2015). Phylogeny and biogeography of Ficedula flycatchers (Aves: Muscicapidae): Novel results from fresh source material. *Molecular Phylogenetics and Evolution, 82*, 87-94.

Moyle, R. G., Oliveros, C. H., Andersen, M. J., Hosner, P. A., Benz, B. W., Manthey, J. D., . . . Faircloth, B. C. (2016). Tectonic collision and uplift of Wallacea triggered the global songbird radiation. *Nature Communications, 7*.

Neall, V. E., & Trewick, S. A. (2008). The age and origin of the Pacific islands: a geological overview. *Philosophical Transactions of the Royal Society B: Biological Sciences, 363*(1508), 3293–3308.

Nyári, Á. S., Benz, B. W., Jønsson, K. A., Fjeldså, J., & Moyle, R. G. (2009). Phylogenetic relationships of fantails (Aves: Rhipiduridae). *Zoologica Scripta, 38*(6), 553–561.

Oliver, L. A., Rittmeyer, E. N., Kraus, F., Richards, S. J., & Austin, C. C. (2013). Phylogeny and phylogeography of *Mantophryne* (Anura: Microhylidae) reveals cryptic diversity in New Guinea. *Molecular Phylogenetics and Evolution, 67*(3), 600–607.

Oliver, P. M., Travers, S. L., Richmond, J. Q., Pikacha, P., & Fisher, R. N. (2017). At the end of the line: independent overwater colonizations of the Solomon Islands by a hyperdiverse trans-Wallacean lizard lineage (*Cyrtodactylus*: Gekkota: Squamata). *Zoological Journal of the Linnean Society, 182*(3), 681–694.

Owens, I. P., Bennett, P. M., & Harvey, P. H. (1999). Species richness among birds: body size, life history, sexual selection or ecology? *Proceedings of the Royal Society of London. Series B: Biological Sciences, 266*(1422), 933-939.

Pedersen, M. P., Irestedt, M., Joseph, L., Rahbek, C., & Jønsson, K. A. (2018). Phylogeography of a 'great speciator'(Aves: *Edolisoma tenuirostre*) reveals complex dispersal and diversification dynamics across the Indo-Pacific. *Journal of Biogeography, 45*(4), 826–837.

Peñalba, J. V., Joseph, L., & Moritz, C. (2019). Current geography masks dynamic history of gene flow during speciation in northern Australian birds. *Molecular Ecology, 28*(3), 630-643.

Pepper, M., Hamilton, D. G., Merkling, T., Svedin, N., Cser, B., Catullo, R. A., . . . Keogh, J. S. (2017). Phylogeographic structure across one of the largest intact tropical savannahs: Molecular and morphological analysis of Australia's iconic frilled lizard Chlamydosaurus kingii. *Molecular Phylogenetics and Evolution, 106*, 217-227.

Petterson, M., Babbs, T., Neal, C., Mahoney, J., Saunders, A., Duncan, R., . . . Mahoa, H. (1999). Geological–tectonic framework of Solomon Islands, SW Pacific: crustal accretion and growth within an intra-oceanic setting. *Tectonophysics, 301*(1-2), 35-60.

Phillimore, A. B., Freckleton, R. P., Orme, C. D. L., & Owens, I. P. (2006). Ecology predicts large-scale patterns of phylogenetic diversification in birds. *The American Naturalist, 168*(2), 220-229.

Polhemus, D. A., Englund, R. A., & Allen, G. R. (2004). *Freshwater Biotas of New Guinea and Nearby Islands: an Analysis of Endemism, Richness, and Threats*: Bishop Museum Technical Report 31.

Pratt, H. D. (2010). Revisiting species and subspecies of island birds for a better assessment of biodiversity. *Ornithological Monographs, 67*(1), 78–89.

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics, 155*(2), 945–959.

Reaz, R., Bayzid, M. S., & Rahman, M. S. (2014). Accurate phylogenetic tree reconstruction from quartets: A heuristic approach. *PloS one, 9*(8), e104008.

Rees, D. J., Emerson, B. C., Oromí, P., & Hewitt, G. M. (2001). The diversification of the genus *Nesotes* (Coleoptera: Tenebrionidae) in the Canary Islands: evidence from mtDNA. *Molecular Phylogenetics and Evolution, 21*(2), 321–326.

Riginos, C., Buckley, Y. M., Blomberg, S. P., & Treml, E. A. (2014). Dispersal capacity predicts both population genetic structure and species richness in reef fishes. *The American Naturalist, 184*(1), 52-64.

Ryan, P. G., Klicka, L. B., Barker, K. F., & Burns, K. J. (2013). The origin of finches on Tristan da Cunha and Gough Island, central South Atlantic ocean. *Molecular Phylogenetics and Evolution, 69*(1), 299–305.

Sayyari, E., & Mirarab, S. (2016). Fast coalescent-based computation of local branch support from quartet frequencies. *Molecular Biology and Evolution, 33*(7), 1654-1668.

Shashank, P., Chakravarthy, A., Raju, B., & Bhanu, K. (2014). DNA barcoding reveals the occurrence of cryptic species in host-associated population of *Conogethes punctiferalis* (Lepidoptera: Crambidae). *Applied Entomology and Zoology, 49*(2), 283–295.

Smith, B. T., McCormack, J. E., Cuervo, A. M., Hickerson, M. J., Aleixo, A., Cadena, C. D., . . . Harvey, M. G. (2014). The drivers of tropical speciation. *Nature, 515*(7527), 406.

Patterns of molecular and morphological variation in some Solomon Island land birds, 2, 124 Cong. Rec. 479-493 (2007).

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics, 30*(9), 1312–1313.

Swofford, D. L. (1998). *PAUP*. Phylogenetic analysis using parsimony (* and other methods). Version 4.* Massachusetts: Sinauer Associates.

Talavera, G., & Castresana, J. (2007). Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Systematic Biology, 56*(4), 564-577.

Taylor, B. (2019). *White-bellied Cuckooshrike (Coracina papuensis). In: del Hoyo, J., Elliott, A., Sargatal, J., Christie, D.A. & de Juana, E. (eds.). Handbook of the Birds of the World Alive. Lynx Edicions, Barcelona.* In.

Toews, D. P., & Brelsford, A. (2012). The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology, 21*(16), 3907-3930.

Toon, A., Drew, A., Mason, I. J., Hughes, J. M., & Joseph, L. (2017). Relationships of the New Guinean subspecies, Gymnorhina tibicen papuana, of the Australian Magpie: an assessment from DNA sequence data. *Emu-Austral Ornithology, 117*(4), 305-315.

Tu, N., Yang, M., Liang, D., & Zhang, P. (2018). A large-scale phylogeny of Microhylidae inferred from a combined dataset of 121 genes and 427 taxa. *Molecular Phylogenetics and Evolution, 126*, 85–91.

Uy, J. A. C., Cooper, E. A., & Chaves, J. A. (2019). Convergent melanism in populations of a Solomon Island flycatcher is mediated by unique genetic mechanisms. *Emu-Austral Ornithology*, 1-9.

Uy, J. A. C., Cooper, E. A., Cutie, S., Concannon, M. R., Poelstra, J. W., Moyle, R. G., & Filardi, C. E. (2016). Mutations in different pigmentation genes are associated with parallel melanism in island flycatchers. *Proceedings of the Royal Society B, 283*(1834), 20160731.

Uy, J. A. C., Moyle, R. G., Filardi, C. E., & Cheviron, Z. A. (2009). Difference in plumage color used in species recognition between incipient species is linked to a single amino acid substitution in the melanocortin-1 receptor. *American Naturalist, 174*(2), 244–254.

Uy, J. A. C., & Vargas-Castro, L. E. (2015). Island size predicts the frequency of melanic birds in the color-polymorphic flycatcher *Monarcha castaneiventris* of the Solomon Islands. *The Auk, 132*(4), 787–794.

Wallace, A. (1881). *Island Life*. New York: Harper & Brothers.

Weeks, B. C., & Claramunt, S. (2014). Dispersal has inhibited avian diversification in Australasian archipelagoes. *Proc Biol Sci, 281*(1791), 20141257. doi:10.1098/rspb.2014.1257

Weidemann, D. E. (2015). The Roles of Geographic Isolation and Gene Flow in the Diversification of the Rufous Fantail *Rhipidura rufifrons*. *Open Access Theses*(587).

Williams, D. J., O'Shea, M., Daguerre, R. L., Pook, C. E., Wuester, W., Hayden, C. J., . . . Winkel, K. D. (2008). Origin of the eastern brownsnake, Pseudonaja textilis (Duméril, Bibron and Duméril)(Serpentes: Elapidae: Hydrophiinae) in New Guinea: evidence of multiple dispersals from Australia, and comments on the status of Pseudonaja textilis pughi Hoser 2003. *Zootaxa, 1703*(1), 47-61.

Wright, S. (1931). Evolution in Mendelian populations. *Genetics, 16*(2), 97.

Zink, R. M. (2004). The role of subspecies in obscuring avian biological diversity and misleading conservation policy. *Proc R Soc Lond B, 271*(1539), 561-564.

## Appendix I

*Rhipidura rufifrons* DAPC results. a) Bayesian Information Criterion indicated the most likely number of genetic clusters for *R. rufifrons* was 7. b) The Solomon Island samples were re-analyzed separately, and BIC again supported a model of 3 genetic clusters within the Solomon Islands.