

7-1-2012

# Diversification in Andean tit-tyrants (Aves, Tyrannidae)

Shane DuBay

Follow this and additional works at: [https://digitalrepository.unm.edu/biol\\_etds](https://digitalrepository.unm.edu/biol_etds)

---

## Recommended Citation

DuBay, Shane. "Diversification in Andean tit-tyrants (Aves, Tyrannidae)." (2012). [https://digitalrepository.unm.edu/biol\\_etds/28](https://digitalrepository.unm.edu/biol_etds/28)

This Thesis is brought to you for free and open access by the Electronic Theses and Dissertations at UNM Digital Repository. It has been accepted for inclusion in Biology ETDs by an authorized administrator of UNM Digital Repository. For more information, please contact [disc@unm.edu](mailto:disc@unm.edu).

Shane G. DuBay  
*Candidate*

Biology  
*Department*

This thesis is approved, and it is acceptable in quality and form for publication:

*Approved by the Thesis Committee:*

Christopher C. Witt, Chairperson

Joseph A. Cook

Steven Poe

**Diversification in Andean tit-tyrants (Aves, Tyrannidae)**

by

**SHANE G. DUBAY**

B.S., Biology, Mercer University, 2007

THESIS

Submitted in Partial Fulfillment of  
the Requirements for the Degree of

**Master of Science**  
**Biology**

The University of New Mexico  
Albuquerque, New Mexico

**July 2012**

## ACKNOWLEDGEMENTS

I thank my advisor and mentor, Dr. Christopher Witt, for providing me with an opportunity to continue my education at the University of New Mexico and always pushing me to be a better scientist. I also thank my thesis committee members, Drs. Joseph Cook and Steven Poe, for providing me with valuable advice during my time at UNM. I learned a great deal from Andrew Johnson who was effectively a second advisor and an integral part of my graduate education.

I thank Phred Benham, Elizabeth Beckman, Natalie Wright, Spencer Galen, Sabrina McNew, Michael Lelevier, Jonathan Schmitt, Ashely Smiley, and Matthew Baumann for helpful discussion and critique. I thank Phred Benham, Natalie Wright, Matthew Jones, Emil Bautista Obispo, Homan Castillo, Andrew Johnson, Jonathan Schmitt, Donna Schmitt, Glenn Seeholzer, and Centro de Ornitología y Biodiversidad (CORBIDI; Thomas Valqui, Emil Bautista Obispo, Willy Nanez) for fieldwork and logistical support to obtain key samples for this study. Lastly, I thank John Kavanaugh, Tye Greene, and Estefany Giehm for keeping me sane during my time at UNM, especially during my final months.

## **Diversification in Andean tit-tyrants (Aves, Tyrannidae)**

by

**Shane G. DuBay**

**B.S., Biology, Mercer University, 2007**

**M.S., Biology, University of New Mexico, 2012**

### **ABSTRACT**

In my thesis I investigate evolution and diversification of montane birds that inhabit dramatic environmental and latitudinal gradients, specifically within the genus *Anairetes* (Aves, Tyrannidae). I use phylogenetic, population genetic, and physiological methods to examine patterns of diversification across environmental gradients. In the first chapter, I infer the phylogeny of *Anairetes* tit-tyrants to provide an essential phylogenetic framework to ask subsequent questions of biogeography and diversification within the group. In the second chapter, I investigated the role of differential adaptation to altitude in promoting diversification and maintaining species limits between *A. reguloides* and *A. nigrocristatus*.

The phylogeny of the flycatcher genus *Anairetes* was previously inferred using short fragments of mitochondrial DNA and parsimony and distance-based methods. The resulting topology spurred taxonomic revision and influenced understanding of Andean biogeography. In the first chapter, I revisit the phylogeny of *Anairetes* tit-tyrants using more mtDNA characters, seven unlinked loci (three mitochondrial genes, six nuclear loci), more closely related outgroup taxa, partitioned Bayesian analyses, and two

coalescent species-tree approaches (Bayesian estimation of species trees, BEST; Bayesian evolutionary analysis by sampling trees, \*BEAST). Of these improvements in data and analyses, the fourfold increase in mtDNA characters was both necessary and sufficient to incur a major shift in the topology and near-complete resolution. The species-tree analyses, while theoretically preferable to concatenation or single gene approaches, yielded topologies that were compatible with mtDNA but with weaker statistical resolution at nodes. Previous results that led to taxonomic and biogeographic reappraisal were refuted, and my results support the resurrection of the genus *Uromyias* as the sister clade to *Anairetes*. The sister relationship between these two genera corresponds to an ecological dichotomy between a depauperate humid cloud forest clade and a diverse dry-tolerant clade that has diversified along the latitudinal axis of the Andes. Species-tree and the concatenation approaches each reaffirm the use of mtDNA to provide phylogenetic signal for avian phylogenies at the species and subspecies level. This result is due in part to the abundance of informative characters in mtDNA, and in part to its lower effective population size that allows it to track the species tree.

Local environmental pressures can drive the evolution of adaptive traits that confer a fitness advantage to an organism under local conditions. In the second chapter I investigate the role of differential physiological adaptation to altitude between sister-species: the elevationally widespread *A. reguloides* and the high-elevation restricted *A. nigrocristatus*. I measure the physiological response of each species to low ambient partial-pressure of oxygen at high elevations. At high elevation, *A. reguloides* shows evidence of hypoxic stress while *A. nigrocristatus* shows evidence of hypoxia resistance. I further quantify the phenotypic and genetic cline shape and the rate and direction of

gene flow between the two species across a narrow contact zone where hybridization occurs at middle elevations. Phenotypic and genetic clines show a dramatic shift from *A. regulooides* to *A. nigrocristatus* across the 212 km elevational transect. Coalescent-based Isolation with Migration (IMa2) analysis suggests effectively zero introgression between parental populations. Upon secondary contact, the two species segregate elevationally. Physiological data, phenotypic and genetic cline shapes, and introgression patterns suggest that maladapted *A. regulooides* alleles are selected against at high elevations in the presence of hypoxia resistant *A. nigrocristatus* alleles. Differential local adaptation is associated with restricted gene flow and essential reproductive isolation, despite at least limited reproductive compatibility.

## Table of Contents

<b>List of Tables.....</b>	<b>ix</b>
<b>List of Figures.....</b>	<b>x</b>
<b>Introduction.....</b>	<b>1</b>
Literature cited.....	4
<b>Chapter 1: An improved phylogeny of Andean tit-tyrants (Aves, Tyrannidae): More characters trump sophisticated analyses.....</b>	<b>7</b>
<b>1. Introduction.....</b>	<b>7</b>
<b>2. Materials and methods.....</b>	<b>10</b>
2.1 Taxonomic and genomic sampling.....	10
2.2 DNA extraction, PCR, sequencing and alignment.....	11
2.3 Data partitions and model selection.....	12
2.4 Gene tree analyses.....	12
2.5 Concatenated analyses.....	13
2.6 Species tree analyses.....	14
<b>3. Results.....</b>	<b>15</b>
3.1 Maximum Parsimony and Bayesian Inference.....	16
3.2 Species tree analysis.....	17
<b>4. Discussion.....</b>	<b>18</b>
4.1 Technological and analytical advances.....	18
4.1.1 Refined sequencing technologies and primer availability.....	18
4.1.2 Partitioning and model-based analyses.....	18
4.1.3 Appropriate outgroup selection.....	19
4.1.4 Coalescent-based analyses of multiple loci.....	20
4.2 Phylogeny of <i>Anairetes</i> .....	22
4.2.1 Taxonomy and topology.....	22
4.2.2 Biogeography.....	24
<b>5. Conclusions.....</b>	<b>26</b>
<b>Acknowledgements.....</b>	<b>28</b>
<b>Literature cited.....</b>	<b>28</b>
<b>Tables.....</b>	<b>38</b>

<b>Figures.....</b>	<b>43</b>
<b>Chapter 2: Diversification by differential adaptation to altitude in Andean tit-tyrants (Aves, Tyrannidae)...</b>	<b>49</b>
<b>1. Introduction.....</b>	<b>49</b>
<b>2. Materials and methods.....</b>	<b>52</b>
2.1 Respiratory physiology.....	52
2.2 Contact zone analysis.....	53
2.2.1 Phenotypic structure.....	54
2.2.2 Genetic structure.....	56
2.2.2.1 DNA extraction, PCR, and sequencing.....	56
2.2.2.2 Population differentiation and cline analysis.....	56
2.2.2.3 Gene flow between <i>A. reguloides</i> and <i>A. nigrocristatus</i> .....	58
<b>3. Results.....</b>	<b>60</b>
3.1 Respiratory physiology.....	60
3.2 Contact zone analysis.....	60
3.2.1 Phenotypic structure.....	60
3.2.2 Genetic structure.....	61
3.3 Evidence of hybridization.....	65
<b>4. Discussion.....</b>	<b>65</b>
4.1 Differential local adaptation.....	66
4.2 Restricted gene flow.....	69
4.3 Mechanism for elevational replacement.....	70
<b>Acknowledgements.....</b>	<b>72</b>
<b>Literature cited.....</b>	<b>72</b>
<b>Tables.....</b>	<b>80</b>
<b>Figures.....</b>	<b>81</b>
<b>Conclusions.....</b>	<b>88</b>

## List of Tables

Table 1. Museum-vouchered samples included in <i>Anairetes</i> phylogeny.....	38
Table 2. Loci included in <i>Anairetes</i> phylogeny and primers used for PCR amplification and sequencing reactions.....	40
Table 3. Summary of sequence variation among loci and substitution models recommended by MrModeltest under AIC, implemented in Bayesian gene tree and species tree analyses.....	42
Table 4. Estimated parameters in the two-population IMA2 analysis.....	80

## List of Figures

Figure 1: Distributions of <i>Anairetes</i> species overlaying a map of mean annual precipitation based on ~1950-2000 (Worldclim v.1.3, 2004).....	43
Figure 2: Gene tree topologies for each nuclear locus obtained from Bayesian Inference and Maximum Parsimony analyses.....	44
Figure 3: Mitochondrial genome topology and concatenated nuclear topology obtained from Bayesian Inference and Maximum Parsimony analyses.....	45
Figure 4: Species-tree topologies recovered from BEST and *BEAST.....	46
Figure 5: Net species and subspecies diversification between the dry-tolerant <i>Anairetes</i> clade and humid <i>Uromyias</i> clade.....	47
Figure 6: Study transect across contact zone of <i>A. reguloides</i> and <i>A. nigrocristatus</i> .....	81
Figure 7: Blood-oxygen carrying capacity parameters for <i>A. reguloides</i> and <i>A. nigrocristatus</i> across elevations.....	82
Figure 8: Morphometric clines across the contact zone of <i>A. reguloides</i> and <i>A. nigrocristatus</i> .....	83
Figure 9: Haplotype networks for mtDNA, Myo2, IRF2, and MUSK.....	84
Figure 10: Bayesian probability of species assignment in Structure across contact zone of <i>A. reguloides</i> and <i>A. nigrocristatus</i> for three nuclear loci.....	85
Figure 11: Genotype clines across the contact zone of <i>A. reguloides</i> and <i>A. nigrocristatus</i> .....	86
Figure 12: Posterior densities of inter-population migration rates from IMa2.....	87

## Introduction

My Master's thesis has been driven by my central scientific question: how do interactions between organisms and their environments affect biodiversity? This question has two components: 1) historical evolutionary processes that have produced the tree of life and 2) current ecological processes that facilitate persistence of biotic lineages in an ecosystem. Investigating the role of differential adaptation on current ecological processes can provide powerful insight into the origins and maintenance of diversity. My thesis investigates how an organism's genetic history dictates its use of the landscape and its interactions with other organisms. More specifically, my thesis examines how local adaptation across elevational gradients affects patterns of diversification and species limits.

Montane regions display conspicuous environmental heterogeneity over small geographical distance. For example, the partial pressure of oxygen ( $PO_2$ ) at 4000m is 40% less than that of sea-level (Beall, 2007). In the Andes of South America this dramatic decrease in ambient  $PO_2$  can occur over a distance of less than 75 km. Other abiotic pressures that change as elevation increases include: a decrease in temperature, an increase in desiccation, and an increase in atmospheric radiation (Monge and Leon-Velarde, 1991; Rezende et al., 2005). The evolution of adaptive traits for hypoxia resistance, cold tolerance, or desiccation resistance is thus beneficial to organisms that inhabit upper montane regions, regardless of the trait's fitness consequence at lower elevations. Genetic adaptation to a particular region of a contiguous environmental gradient must occur despite antagonistic pleiotropy, whereby locally advantageous alleles are deleterious outside of local conditions (Williams, 1966; Kawecki and Elbert, 2004). Lineages often span heterogeneous environments and can show phenotypic and genetic differentiation as a result of contrasting selective pressures imposed by disparate

environmental conditions (Felsenstein, 1976; Hedrick et al., 1976; Hedrick, 1986; Schneider et al. 1999; Cheviron and Brumfield, 2009). Gene flow, however, counteracts local differentiation by homogenizing populations. In the face of gene flow, differentiation can be maintained only by strong selection in favor of locally adapted alleles. Thus, the potential for local adaptation and divergence is highest when selection is strong and gene flow is low.

In my thesis I investigate the role of differential local adaptation to high elevation, specifically adaptation to low ambient PO<sub>2</sub>, in promoting diversification and maintaining species limits and distributional limits between sister-species of tit-tyrant flycatchers in the genus *Anairetes* (Aves, Tyrannidae). Eight species are currently recognized within *Anairetes*, all occurring in western South America. Members of the genus primarily inhabit high elevations along the spine of the Andes. In the first chapter of my thesis I revise the phylogeny of *Anairetes* tit-tyrants to provide a robust phylogenetic platform to address subsequent questions of diversification within the group. In the second chapter I investigate the role of differential physiological adaptation to altitude in maintaining species limits between *A. reguloides* and *A. nigrocristatus* upon secondary contact, despite hybridization.

*Anairetes* was the target of one of the first comprehensive molecular phylogenetic studies of an avian genus. The initial phylogeny used short fragments of mitochondrial DNA and parsimony and distance-based methods (Roy et al., 1999). The resulting topology recovered weak internal resolution but nested the genus *Uromyias* within *Anairetes sensu stricto*. The placement of *Uromyias* within *Anairetes* was not statistically supported but provided the impetus for taxonomic authorities to dissolve the genus *Uromyias* (e.g. Remsen et al., 2003; Clements, 2004; Del Hoyo et al., 2004; Schulenberg et al., 2007; Gill and Donsker, 2008). The correct placement of *Uromyias* with respect to *Anairetes sensu stricto* is important for biogeographic

interpretation of Andean diversification patterns. If *Uromyias* arose within *Anairetes* then it suggests that a clade specialized on humid forests evolved from a generalist ancestor tolerant of dry habitats. If *Uromyias* is sister to *Anairetes sensu stricto* then it implies a deeper ecological dichotomy where a dry-tolerant clade has undergone greater morphological, ecological, and lineage diversification than a humid-forest specialist clade. In the first chapter of my thesis I revisit the phylogeny of *Anairetes* using multiple unlinked loci, partitioned model-based analyses, appropriate outgroups, and coalescent-based species-tree methods.

The second chapter of my thesis uses the phylogenetic context of chapter one to ask what happens when latitudinally segregated allotaxa become elevational replacements in secondary contact, a key stage in the diversification of Andean birds. *A. reguloides* spans an enormous elevational distribution on the west slope of the Andes (sea-level to 4200m) while *A. nigrocristatus* is restricted to upper montane regions (3000m to 4200m). In the absence of gene flow from external environments, locally adapted traits should be driven to fixation within a population. I hypothesize that *A. nigrocristatus* will show evidence of high-elevation specialization because it is restricted to cold, hypoxic upper elevations. Thus, *A. reguloides* should show evidence of hypoxic stress at high elevation. Gene flow with lowland conspecifics would prevent high-elevation adaptation from evolving or persisting in *A. reguloides*. I measured blood-oxygen carrying capacity parameters that affect oxygen delivery to respiring tissue and respond to changes in environmental  $PO_2$  to address differential physiological adaptation between the two species. I use these parameters as an inverse index of respiratory performance because they indicate hypoxic stress as a function of ambient  $PO_2$ , and possibly also temperature.

During my study I discovered a hybrid zone at middle elevations between *A. reguloides* and *A. nigrocristatus* where phenotypically and genetically intermediate birds occur. In my

second chapter I compare the physiological data to the phenotypic and genetic structure across an elevational transect that spans the contact zone of the two parental populations. I examine introgression by quantifying morphological and genetic cline shapes and calculating the rate and direction of gene flow along the transect. I hypothesize that differential adaptation to altitude promotes elevational segregation between *A. reguloides* and *A. nigrocristatus*, despite at least limited hybridization. *A. reguloides* alleles are likely maladapted to high elevations, providing the basis for competitive displacement by *A. nigrocristatus* at high elevations.

My thesis combines phylogenetic, population genetic, and physiological data to investigate the role of historical evolutionary selection on shaping current ecological processes and species dynamics. Lineages are often subjected to disparate environmental pressures that can shape the trajectory of adaptive trait evolution. Montane regions show contrasting selective pressures over small geographical distance, which make these regions ideal for studies of adaptive evolution and speciation. The goal of my thesis is to provide further understanding of diversification processes in montane regions and across heterogeneous landscapes.

### **Literature cited**

Beall, C. 2007. Two routes to functional adaptation: Tibetan and Andean high-altitude natives.

Proceedings of the National Academy of Sciences of the United States of America  
104:8655-8660.

Cheviron, Z., and R. Brumfield. 2009. Migration-selection balance and local adaptation of mitochondrial haplotypes in Rufous-collared Sparrow (*Zonotrichia capensis*) along an elevational gradient. *Evolution* 63:1593-1605.

- Clements, J., 2004. The Clements Checklist of the Birds of the World (5 ed.). Ithaca, New York, Cornell University Press.
- Del Hoyo, J., Elliot, A., Sargatal, J., Christie, D. A. 2004. Handbook of the birds of the world, Vol. 9. Cotingas to Pipits and Wagtails. Lynx, Barcelona.
- Felsenstein, J. 1976. Theoretical population genetics of variable selection and migration. Annual Review of Genetics 10:253-280.
- Gill, F., Donsker, D., 2008. IOC World Bird Names (version 1.5).  
<<http://www.worldbirdnames.org>>
- Hedrick, P. 1986. Genetic polymorphism in heterogeneous environments- a decade later. Annual Review of Ecology and Systematics 17:535-566.
- Hedrick, P., M. Ginevan, and E. Ewing. 1976. Genetic polymorphism in heterogeneous environments. Annual Review of Ecology and Systematics 7:1-32.
- Kawecki, T., and D. Ebert. 2004. Conceptual issues in local adaptation. Ecology Letters 7:1225-1241.
- Monge, C., and F. Leonvelarde. 1991. Physiological adaptation to high-altitude oxygen transport in mammals and birds. Physiological Reviews 71:1135-1172.
- Remsen, J. V., Jr., Cadena, C. D., Jaramillo, A., Nores, M., Pacheco, J. F., Pérez-Emán, J., Robbins, M. B., Stiles, F. G., Stotz, D. F., Zimmer, K. J., 2003. A classification of the bird species of South America. American Ornithologists' Union.  
<<http://www.museum.lsu.edu/~Remsen/SACCBaseline.html>>
- Rezende, E., F. Gomes, C. Ghalambor, G. Russell, and M. Chappell. 2005. An evolutionary frame of work to study physiological adaptation to high altitudes. Revista Chilena De Historia Natural 78:323-336.

- Roy, M. S., Torres-Mura, J. C., Hertel, F., 1999. Molecular phylogeny and evolutionary history of the tit-tyrants (Aves: Tyrannidae). *Molecular Phylogenetics and Evolution* 11, 67-76.
- Schneider, C., T. Smith, B. Larison, and C. Moritz. 1999. A test of alternative models of diversification in tropical rainforests: Ecological gradients vs. rainforest refugia. *Proceedings of the National Academy of Sciences of the United States of America* 96:13869-13873.
- Schulenberg, T. S., Stotz, D. F., Lane, D. F., O'Neill, J. P., Parker, T. A., III. 2007. *Birds of Peru*. Princeton University Press, New Jersey.
- Williams, G. C. 1966. *Adaptation and Natural Selection*. Princeton University Press, Princeton, New Jersey.

## Chapter 1: An improved phylogeny of Andean tit-tyrants (Aves, Tyrannidae):

### More characters trump sophisticated analyses

Published in: DuBay, S. G., and C. C. Witt. 2012. An improved phylogeny of Andean tit-tyrants: More characters trump sophisticated analyses. *Molecular Phylogenetics and Evolution*.

<http://dx.doi.org/10.1016/j.ympev.2012.04.002>

#### 1. Introduction

The rise of DNA sequence-based phylogenies over the past two decades has fostered steady improvement in our understanding of avian biogeography and prompted a shift towards a phylogenetic classification of birds. While avian phylogeneticists proceed towards comprehensive taxon sampling there is ample need to revisit phylogenetic hypotheses whose influence may not have been justified by empirical support. Roy et al. (1999) provided one of the first comprehensive species-level molecular phylogenies of an avian genus, *Anairetes* (tit-tyrant flycatchers). The resulting topology spurred a taxonomic revision and influenced subsequent researchers with respect to biogeography and avian diversification patterns in the Andes [e.g. (Moritz et al., 2000; Webb and Gaston, 2003; Dingle et al., 2006; Boyle and Conway, 2007)]. In the ensuing decade, numerous technological and analytical advances have improved prospects for accurate estimation of phylogenies, including: (1) primer sequences and refined sequencing technologies to obtain long sequences from multiple unlinked loci (Sehgal and Lovette, 2003; Kimball et al., 2008); (2) increased computational power for tree search, parameter estimation, and bootstrapping (Brownstone and Valletta, 2001); (3) data partitioning and posterior probability estimation using model-based Bayesian analysis (Huelsenbeck and Ronquist, 2001;

Huelsenbeck and Crandall, 1997; Huelsenbeck and Rannala, 1997; Brandley et al., 2005); (4) higher-level phylogenetic structure for a substantial portion of extant birds that aids in the selection of appropriate outgroups (e.g. Ohlson et al., 2008; Tello et al., 2009); and (5) coalescent-based analyses of multi-locus datasets (Degnan and Salter, 2005; Drummond and Rambaut, 2007; Liu and Pearl, 2007; Liu, 2008;). Importantly, the latter advance has initiated a shift from multi-locus concatenation to a gene-tree coalescent approach (Edwards, 2009).

Eight species and 17 subspecies are currently recognized in the genus *Anairetes* (Dickinson, 2003). The inclusion of *A. agraphia* and *A. agilis* and subspecies therein (hereforward called *Uromyias*) has been disputed since the early twentieth century. Hellmayr (1927) and Lanyon (1988) recognized *A. agraphia* and *A. agilis* as a distinct genus, *Uromyias*, while Smith (1971) and Traylor (1977) recognized the two species as members of *Anairetes*. Roy et al. (1999) addressed the phylogenetic placement of *Uromyias* using short mitochondrial DNA fragments (totaling 632 base pairs of ND2 and Cyt b) and techniques available at the time. “*Uromyias agilis*” was recovered nested within *Anairetes sensu stricto* based on parsimony analysis of 342 bp of ND2 sequence and Kimura 2-parameter distances of the combined mtDNA dataset with 54% and <50% bootstrap support, respectively. Neither topology recovered *Uromyias* definitively outside *Anairetes*. Roy et al. (1999) were conservative in interpreting their results because of low resolution at basal nodes, although they stated plainly that “molecular data do not support a monophyletic arrangement of *Anairetes* relative to *Uromyias*.” Subsequently, the genus *Uromyias* was merged back to *Anairetes* by major taxonomic authorities (Remsen et al., 2003; Clements, 2004; Del Hoyo et al., 2004; Schulenberg et al., 2007; Gill and Donsker, 2008).

The phylogenetic placement of *Uromyias* (*A. agraphia* and *A. agilis*) has implications for our understanding of the biogeography of the Andes. These two species are the only members of the *Anairetes* group that are restricted to extreme humid habitats. The other continental *Anairetes* species are more tolerant of arid conditions and all occur at high elevations in rain-shadow valleys and on west-facing slopes that are subject to at least seasonal aridity (Fig. 1). These dry-tolerant *Anairetes* species are all sympatric or syntopic with other *Anairetes* species in parts of their distributions, with as many as three species occurring together (e.g. *A. parulus*, *A. reguloides*, and *A. flavirostris* in western Peru). In contrast, the members of *Uromyias* are exclusively allopatric or parapatric with other members of the group. *Uromyias* is the only subclade whose distribution is not primarily south of the equator (*A. parulus aequatorialis* extends to southern Colombia and is the only other taxon in the genus to occur north of the equator). This pattern is apparent in other avian taxa with dry-tolerant montane genera becoming scarce north of the equator, reflecting the humidity gradient along the latitudinal axis of the Andes (e.g. *Asthenes*, *Muscisaxicola*, *Phrygilus*, *Oreotrochilus*, *Geositta*, *Leptasthenura*). If *Uromyias* is nested within *Anairetes sensu stricto*, it implies that the humid cloud-forest specialist clade evolved from a dry-tolerant habitat generalist ancestor. Conversely, if *Uromyias* represents a sister clade it would imply an older ecological dichotomy and a subsequent difference in net diversification, with the 13 dry-tolerant, southern lineages having undergone more morphological, ecological, and lineage diversification than the four humid-restricted taxa.

In this study, we estimate the phylogeny of the genus *Anairetes* using seven loci, 6407 base pairs, partitioned Bayesian analysis, species-tree methods, and appropriate outgroups. By revisiting the Roy et al. (1999) study, we examine the contributions of each technological and analytical advance to changes in the topology and its nodal resolution. In particular, we compare

the performance of multiple nuclear loci versus mitochondrial DNA and species-tree methods versus concatenation. We aim to confidently place *Uromyias* with regard to *Anairetes sensu stricto*. Accurate placement of *Uromyias* will provide essential framework to better understand the role of dry and humid habitat pressures in shaping the diversification patterns within the group.

## 2. Materials and methods

### 2.1 Taxonomic and genomic sampling

We sampled the eight currently recognized species in the genus *Anairetes*, including 14 of the 17 recognized subspecies (Table 1; Dickinson, 2003). We sampled *A. nigrocristatus* and *A. reguloides albiventris* outside of a zone of potential introgression in Ancash, Peru. We were unable to obtain samples of *A. agraphia plengei*, *A. agraphia squamigera*, or endangered *A. alpinus alpinus*. Samples of *A. fernandezianus*, endemic to Robinson Crusoe Island, Chile, were not available; however, mtDNA sequences were available on Genbank (Roy et al., 1999). We selected five outgroup taxa on the basis of recent phylogenies of the Tyrannidae (Ohlson et al., 2008; Tello et al., 2009), including *Culicivora caudacuta*, *Mecocerculus leucophrys*, *Polystictus pectoralis*, *Pseudocolopteryx sclateri*, and *Serpophaga munda*. These five taxa and *Anairetes* comprise a clade within the Elaeniinae assemblage of Tyrannidae with *Mecocerculus leucophrys* positioned basally.

For the majority of samples we obtained DNA sequences for three protein-coding mitochondrial genes [NADH dehydrogenase subunit 2 (ND2), subunit 3 (ND3), cytochrome b (Cyt b)], three autosomal nuclear intron loci [interferon regulatory factor 2 (IRF2), myoglobin intron 2 (Myo2), period homolog 2 (PER2)], two autosomal nuclear exon loci [brain-derived neurotrophic factor (BDNF), nerve growth factor (NGF)], and a sex-linked nuclear intron locus

[muscle, skeletal, receptor tyrosine kinase (MUSK); (Table 1)]. Nuclear markers were chosen largely based on their location in the chicken genome; we chose unlinked markers on different chromosomes that are known to have different evolutionary rates, selection pressures, and population sizes. Four sequences were obtained from GenBank (*A. agilis*, ND2; *A. fernandezianus*, ND2 and Cyt b; *Serpophaga munda*, Myo2). Sequences for each marker came from the same individual for all other taxa (Table 1). We obtained 112 of 126 possible sequences of the nine genes from the 14 ingroup taxa and 44 of 45 sequences from the five outgroup taxa. The missing sequences were: Cyt b for *Polystictus pectoralis*, PER2 for *A. parulus patagonicus*, ND3 for *A. fernandezianus*, and nuclear sequences for *A. agilis* and *A. fernandezianus* (Table 1). The only *A. agilis* samples available to us were degraded, and I was thus unable to amplify and sequence nuclear loci. Fortunately, the sister relationship of *A. agilis* and *A. agraphia* is uncontroversial and is consistent with mtDNA and all previous taxonomic treatments.

## 2.2 DNA extraction, PCR, sequencing and alignment

I extracted total DNA from frozen and ethanol-preserved skeletal muscle using the DNeasy Tissue Kit (Qiagen, Valencia, CA). Primers and conditions for PCR amplification for eight of the nine markers were obtained from the literature (see Table 2 for a list of primers used and source). I designed *Anairetes*-specific primers for MUSK, Myo2, and PER2 using Primer 3v.0.4.0 (Rozen and Skaletsky, 2000) from successful *Anairetes sp.* sequences obtained from the literature-specified primers. I used PCR and sequencing protocols described by Johnson et al. (2011). I did not detect evidence of pseudogenes in the mtDNA data. Unambiguous double-peaks in the nuDNA of equal height at the same nucleotide position were coded as ambiguous. I aligned sequences with MUSCLE 3.7 (Edgar, 2004) and I inspected alignments and assigned codon positions using MacClade 4.08 OS X (Maddison and Maddison, 2005).

### 2.3 Data partitions and model selection

I conducted preliminary Maximum Parsimony and Bayesian inference analyses separately for each of the mtDNA genes. The mtDNA gene trees showed 100% concordance at the interspecific level and were thus concatenated for further analyses. For Bayesian phylogenetic analyses I identified the best-fitting model for each of the seven loci, with the concatenated mtDNA partitioned by codon position (mtDNA codon pos.1, mtDNA codon pos.2, mtDNA codon pos. 3, BDNF, NGF, IRF2, Myo2, PER2, MUSK), using MrModeltest v2 (Nylander, 2004) with the Akaike information criterion (AIC, Akaike, 1974). Table 3 shows the preferred model for each locus. The concatenated mtDNA was partitioned by codon position to account for different evolutionary rates and selective pressures among position (Table 3; Bull et al., 1993). I did not partition BDNF and NGF by codon position although they are protein-coding genes because limited to no variation was observed in these genes when initially partitioned. For example, BDNF codon position two and NGF codon position one were homogeneous across all study taxa. Non-coding nuclear intron loci were also treated as individual partitions.

### 2.4 Gene tree analyses

I estimated the mitochondrial topology and each nuclear gene tree with Maximum Parsimony (MP) and Bayesian inference (BI). The concatenated mtDNA and individual nuclear genes datasets were run independently using the two phylogenetic methods to investigate congruence among loci. I conducted MP analyses in PAUP\* 4.0b10 (Swofford, 2002), implementing the branch-and-bound search algorithm with characters equally weighted. I assessed nodal support with 1000 bootstrap pseudoreplicates under a heuristic search, tree bisection-reconnection (TBR) branch swapping, and 100 random sequence additions (Felsenstein, 1985). Bayesian analyses were conducted in MrBayes 3.2.1 (Ronquist and

Huelsenbeck, 2003) consisting of two replicated runs for each locus with four MCMC chains with default heating (1 cold, 3 heated). Each analysis ran for 15,000,000 generations, sampling every 1000 generations. I assessed likelihood stabilization and convergence between runs using Tracer v.1.5 (Rambaut and Drummond, 2009). I discarded the first 25% of trees as burn-in although I achieved split frequencies  $<0.005$  and stationary likelihood values much sooner.

Mitochondrial protein-coding genes show evidence of episodic positive selection or relaxed purifying selection in conjunction with changing thermal environments (e.g. Gering et al., 2009). *Anairetes* spans dramatic gradients in elevation and temperature; accordingly, I repeated the mtDNA phylogenetic analysis using only the third codon positions to assess the potential sensitivity of the topology to selection.

In addition to the independent gene tree analyses, I conducted MP and BI analyses of the Roy et al. (1999) mitochondrial dataset, replacing their outgroup, *Stigmatura*, with the more closely related, *Serpophaga munda* (Ohlson et al., 2008; Tello et al., 2009), to test for the effects of appropriate outgroup selection on tree topology.

## 2.5 Concatenated analyses

Individual genes alone may lack sufficient variation or signal to provide resolution or strong nodal support for bipartitions. Concatenating individual genes can increase phylogenetic signal (if complimentary and not conflicting) and recover bipartitions or increase nodal support not evident in individual gene analyses (Barrett et al., 1991; Chippindale and Wiens, 1994). I conducted MP and BI analyses on two concatenated matrices (total dataset and nuclear dataset alone) under the conditions described in 2.4 to investigate the presence or absence of this emergent signal phenomenon within our dataset. I excluded *A. agilis* and *A. fernandezianus* from the total concatenated dataset because they lacked nuclear sequences.

In addition to MP and BI analyses, I also conducted preliminary analyses of individual loci and concatenated datasets using maximum likelihood methods. The resulting topologies were highly consistent with our parsimony and Bayesian analyses regarding the respective dataset.

## 2.6 Species tree analyses

I independently estimated the species tree for the genus *Anairetes* using two methods: BEST 2.3.1 (Liu, 2008) implemented in MrBayes and \*BEAST 1.6.2 (Drummond and Rambaut, 2007). Both methods employ a coalescent model that assumes that any discordance between gene topologies resulted from ancestral polymorphism (i.e. incomplete lineage sorting). BEST reconstructs a species tree from fixed gene trees, while \*BEAST simultaneously co-estimates the species tree and gene trees (Heled and Drummond, 2010). For each method, a single species tree was constructed from the seven independent loci (mtDNA, BDNF, IRF2, PER2, MUSK, Myo2, NGF), implementing the recommended substitution models for each locus obtained from MrModeltest. Only taxa with sequence data for every locus were used.

In BEST, a single model (GTR+G+I) was assigned to the mtDNA locus since I could not subpartition by codon position. I conducted two replicated runs in BEST for 300 million generations, sampling every 100,000 generations with four MCMC chains with default heating (1 cold, 3 heated). I used a flat-prior distribution of population size (inverse gamma distribution,  $\alpha=3$  and  $\beta=.003$ ) and uniform distributions of mutation rate (bounded values 0.5 and 1.5) as suggested by the authors and previous phylogenetic studies of birds (Brumfield et al., 2008; Liu et al., 2008). I unlinked substitution model parameters between partitions, and I specified the mtDNA locus as haploid and the remaining nuclear loci as diploid. Haploid specification in BEST accounts for the one-fourth smaller effective population size of mtDNA to nuDNA as a

consequence of its haploid nature and matrilineal inheritance (Moore, 1995; Liu and Pearl, 2007; Waters et al., 2010).

In \*BEAST, the mtDNA was partitioned by codon position and a single model (GTR+G+I) was implemented for each position. I unlinked substitution model parameters between partitions. Nuclear alleles were not phased because populations were represented by single individuals. I formatted our input file using BEAUti 1.7 included in the BEAST software package. I specified a Yule tree prior and a random starting tree for each locus. A Yule speciation process is most appropriate when comparing relationships between species or when populations are represented in the data by single individuals ([http://beast.bio.ed.ac.uk/Tree\\_priors\\_and\\_dating](http://beast.bio.ed.ac.uk/Tree_priors_and_dating); Patterson et al., 2011). I compared strict vs. relaxed molecular clock models using a likelihood ratio test and I found no significant departure from a strict clock; thus, a strict molecular clock was implemented in final analyses. In the absence of reliable calibration points (i.e. fossil data), I implemented a fixed mean substitution clock rate of 1.0 as suggested by authors (Drummond et al. 2007). Setting a fixed mean substitution rate is appropriate when the goal of analyses is phylogenetic reconstruction and not divergence dating (Drummond et al. 2007). I conducted two replicated runs in \*BEAST for 400 million generations, sampling every 100,000 generations. I used Tracer v.1.5 (Rambaut and Drummond, 2009) to assess likelihood stabilization and convergence for BEST and \*BEAST analyses. I discarded the first 50% of trees as burn-in for the two species-tree methods.

### **3. Results**

#### **3.1 Maximum Parsimony and Bayesian Inference**

There was no strongly supported conflict between MP and BI analyses of the same dataset for all gene trees and concatenated topologies. I consider strong nodal support as non-

parametric bootstrap values >70% and Bayesian posterior probabilities >95%. Nuclear gene trees lacked resolution at the majority of bipartitions (Fig. 2). Myo2 had the highest resolution among nuclear loci, recovering *A. nigrocristatus/reguloides* as sister to *A. alpinus/flavivostriis/parulus* with strong support. BDNF and NGF did not recover any nodes with strong support. At strongly supported nodes individual gene trees recovered each species as monophyletic, with one exception: MUSK placed *A. flavivostriis huancabambae* basal to *A. alpinus*, making *A. flavivostriis* paraphyletic (Fig. 2). This exception was the only strongly supported discordant node among gene trees (Table 4).

The mtDNA Bayesian tree resolved every *Anairetes* interspecific relationship with posterior probabilities >99% (Fig. 3). The Bayesian tree based on concatenation of all genes (not shown) was identical to the mtDNA Bayesian tree in branching structure and almost identical in nodal support values. This result comports with the substantially higher variation in the mtDNA compared to nuDNA (Table 3). The concatenated nuDNA recovered the same ingroup topology as the mtDNA with the exception: *A. nigrocristatus* sister to *A. reguloides albiventris* to the exclusion of *A. r. reguloides* (Fig. 3b). The resulting paraphyly of *A. reguloides* in the concatenated nuDNA tree was not strongly supported by any individual nuclear gene tree. The concatenated nuDNA tree showed increased resolution over individual nuclear gene trees, including increased nodal support values and novel bipartitions not evident in any individual nuclear gene tree (Fig. 3; evidence of emergent signal).

Contrary to the tree obtained by Roy et al. (1999), both our mtDNA tree and concatenated nuDNA tree place *Uromyias* outside the historical genus *Anairetes*. I recovered two clades that were consistent with the Roy et al. (1999) findings: (1) *A. nigrocristatus/reguloides* and (2) *A. alpinus/flavivostriis/fernandezianus/parulus*. However, Roy

et al. (1999) placed *A. alpinus* sister to *A. flavirostris* with 67% bootstrap support (Fig. 3c). In our analysis both mitochondrial and nuclear loci strongly support the positioning of *A. alpinus* as sister to the clade containing *A. flavirostris* and *A. parulus/fernandezianus* (Fig. 3). The topology based on the 3<sup>rd</sup> codon position of mtDNA was the same as that based on the whole mtDNA but with slightly weaker nodal support. This result suggests that potential selection on mitochondrial coding genes is not driving topology.

Using a more appropriate outgroup with the Roy et al. (1999) mitochondrial dataset apparently caused subtle changes in nodal support under MP and BI, but the overall topology was unchanged from the original. Resolution at internal nodes remained weak and *A. agilis* remained within *Anairetes sensu stricto*.

### 3.2 Species tree analysis

The species trees produced using BEST and \*BEAST strongly supported the majority of interspecific relationships recovered by the mtDNA and concatenated nuDNA (posterior probabilities >95%; Fig. 4). Nodal support values were lower in the species-tree topologies than in the mtDNA and concatenated trees. BEST still recovered seven nodes with strong support while \*BEAST recovered six. The species-tree topologies strongly supported the placement of *A. alpinus* sister to *A. flavirostris/parulus* and placed the *A. alpinus/flavirostris/parulus* clade sister to *A. nigrocristatus/reguloides*, corroborating the mtDNA tree and concatenated nuDNA tree, and confirming the monophyly of the core *Anairetes* clade. In the BEST topology, *Uromyias* was not sister to the core *Anairetes* clade, however, it was nested among the outgroup taxa, albeit with weak resolution. I suspect BEST had trouble converging; analysis reached stationary likelihood values, but three of the seven gene trees had split frequencies >0.05 (MUSK= 0.71, Myo2 = 0.72, NGF = 0.56; Hall, 2007). The \*BEAST topology showed no strongly supported

discordance with the mtDNA tree and recovered *Uromyias* as sister to the core *Anairetes* clade with 100% posterior probability (Fig. 3a and 4b).

## 4. Discussion

### 4.1 Technological and analytical advances

#### 4.1.1 Refined sequencing technologies and primer availability

Increased availability of nuclear markers and ease of sequencing allows for the acquisition of long sequences from multiple unlinked loci. Increasing the number of mtDNA characters from those used by Roy et al. (1999) significantly changed tree topology and improved nodal support. *Uromyias* was recovered outside of *Anairetes sensu stricto*, and *A. parulus/fernandezianus* was recovered sister to *A. flavirostris*. The additional mtDNA characters alone led to an increase in the number of interspecific bipartitions with bootstrap values >99% from two to six, out of seven possible. Sequencing multiple unlinked loci provided little increase in resolution in itself (under MP), a result that is likely attributable to limited phylogenetic information in our nuclear loci compared to mtDNA. The maximum p-distance of nuclear loci ranged from 0.000-0.017 within ingroup taxa; the maximum p-distance of the mtDNA was .117, highlighting the greater levels of divergence in the mitochondrial genome and the limited number of informative nuclear sites (Table 3).

#### 4.1.2 Partitioning and model-based analyses

Partitioning data with model-based analyses recovered more bipartitions and increased nodal support in gene trees and concatenated trees alike. BI (partitioned and model-based) recovered 12 of the 14 strongly supported nodes recovered by MP in individual gene trees and the same four strongly supported nodes in the nuclear concatenation topology, but BI also recovered seven and five additional nodes, respectively. In the six individual nuclear gene trees

BI recovered eight strongly supported nodes while MP recovered only three. Parsimony analysis of mtDNA, which had over 44 times the amount of informative variation of any single nuclear locus, recovered 10 of the 12 bipartitions that were strongly supported by the model-based BI analysis of mtDNA, with no conflicts. Model-based analyses had higher resolution at more bipartitions than MP in both cases, but the improvement over MP was more pronounced for nuclear genes, where phylogenetic information was limited. This observation was unanticipated because I expected that model-based analyses would outperform MP with mtDNA data since it is more likely to be saturated and have homoplastic sites. The difference between MP and BI results for nuclear DNA is difficult to reconcile with the observation that MP estimates are effectively maximum likelihood estimates in the absence of homoplasy (Steel and Penny 2000).

#### 4.1.3 Appropriate outgroup selection

Recent higher-level phylogenetic structure of genera within Tyrannidae allowed for the closest known outgroups to be selected for this study (Ohlson et al., 2008; Tello et al., 2009). Roy et al. (1999) chose *Stigmatura napensis* as the outgroup based on the morphological classification of tyrannid flycatchers by Lanyon (1988). Ohlson et al. (2008) and Tello et al. (2009) independently recovered five genera that comprise a clade with *Anairetes* to the exclusion of the genus *Stigmatura*. Our re-analyses of the Roy et al. (1999) dataset with more closely related outgroups produced a tree that was consistent with the topology and resolution of Roy et al. (1999) regardless of phylogenetic method.

#### 4.1.4 Coalescent-based analyses of multiple loci

Our species-tree topologies were highly consistent with our mtDNA topology but had weaker nodal support. Our results confirm previous results and predictions that coalescent-based species-tree methods generally have lower statistical confidence than concatenation methods

(Belfiore et al., 2008; Thomson et al., 2008; Edwards, 2009). Data “swamping” could potentially cause this pattern, whereby the concatenated topology is biased by one or a few loci that provide the majority of phylogenetic signal, resulting in artificially high nodal support and unrealistic branch lengths (Hillis, 1987; Baker et al., 1998; Edwards, 2009). This phenomenon might explain the results of our total concatenated analysis where the recovered topology was identical to the mtDNA topology. Alternatively, species-tree approaches separately consider the phylogenetic signal from each gene, preventing highly variable loci from “swamping” the tree topology (Hillis, 1987; Baker et al., 1998, Edwards, 2009). Even when gene trees recover identical topologies, significant branch length heterogeneity can produce unexpected and incorrect phylogenetic signal when concatenated (Kolaczkowski and Thornton, 2004; Matsen and Steel, 2007). Edwards (2009) suggests that coalescent-based methods provide more realistic nodal support and branch lengths than concatenation by equally weighting the phylogenetic signal of each locus and explicitly accounting for topological heterogeneity among loci.

BEST and \*BEAST take into consideration the one quarter effective population size of mtDNA relative to nuDNA through ploidy specification (Liu and Pearl, 2007; Drummond and Rambaut, 2007). If the mtDNA is not specified as haploid then the mtDNA tree is considered to be just as likely as any nuclear locus to show discordance with the species tree. In BEST, setting small theta values reduces the influence of effective population size on the resulting species tree by increasing the probability for a speciation event to occur (Liu et al., 2008; Leache and Fujita, 2010), but a uniform small theta does not account for the known differences between mitochondrial and nuclear loci in accurately tracking the species topology. I ran preliminary BEST analysis under the conditions described in section 2.5 but treated mtDNA as a diploid locus to test the effect of ploidy specification. The topology was highly consistent with our

depicted BEST tree (Fig. 3) but had one major conflicting species relationship: *A. reguloides albiventris* was sister to *A. nigrocristatus* to the exclusion of *A. r. reguloides*. This was the same relationship recovered in the nuclear concatenated topology. When the mtDNA was not specified as haploid the resulting topology implied, improbably, that (1) the mtDNA gene tree contained a deep coalescence event, (2) the concatenated nuDNA tracked the correct species tree, and (3) that a morphologically well-defined species is paraphyletic. The latter result is most likely incorrect and highlights the primacy of mtDNA in resolving species and subspecies relationships in birds.

It is important to note that BEST had convergence problems with our dataset. Independent runs in BEST strongly corroborate the topology within *Anairetes sensu stricto* but appear to have trouble resolving basal topology. Convergence problems have been documented in empirical studies of Locustellid warblers (Alstrom et al., 2011), *Zea* maize (Cranston et al., 2009), and *Neodiprion* saw flies (Linnen and Farrell, 2008). These previous studies suggest that increasing the size of datasets may require exceptionally long runs in BEST to reach convergence. Waters et al. (2010) reached convergence after 20 million generations with 11 taxa of *Galaxias* fish and four genes by running six MCMC chains (1 cold, 5 at low heat of 0.1; suggested by Beiko et al., 2006). Increasing the number of heated MCMC chains increases mobility while searching tree space and decreases the probability of getting trapped in local optima. In our study, I initially ran BEST with two MCMC chains. The preliminary analyses resulted in poor convergence. For final analyses I increased the number of MCMC chain from two to four (1 cold, 3 heated). I could not run six MCMC chains as suggested by Beiko et al. (2006) because I lacked sufficient computational time and power. Increasing the number of

MCMC chains from two to four helped with convergence, but high split frequency values were still apparent, suggesting that our independent runs had not yet converged.

## 4.2 Phylogeny of *Anairetes*

### 4.2.1 Taxonomy and topology

The placement of *Uromyias* (*A. agraphia* and *A. agilis*) has been debated for decades based on morphological differences and more recently, DNA sequence data (Roy et al., 1999). *A. agraphia* and *A. agilis* were first described as a distinct superspecies within the genus *Anairetes* based on bill and tail morphology (Sclater, 1888; Chapman, 1919). Hellmayr (1927) placed the two species in a new genus, *Uromyias*, based on morphological characters including: a shorter, wider and more depressed bill, more developed rictal bristles, proportionately longer tail, and greater variation between the shortest and longest rectrices. Smith (1971) replaced *A. agraphia* and *A. agilis* to *Anairetes* based on morphological similarity and the premise that ecological differences should not delimit generic boundaries. Traylor (1977) supported dissolving the genus *Uromyias* and argued that the most recently described species, *A. alpinus* (Carriker, 1933), is morphologically intermediate between the two genera. Traylor (1977) argued that the morphological differences between *Uromyias* and *Anairetes* “do not seem of great importance in an otherwise closely related group”. Lanyon (1988) supported the validity of the genus *Uromyias* based on cranial morphology; specifically, *Uromyias* has a fully ossified nasal septum and lacks posterior forking in the trabecular plate. Further, Lanyon (1988) argued that posterior forking of the trabecular plate suggests monophyly for an *Anairetes/Serpophaga* clade, excluding *Uromyias*. Roy et al. (1999) provided the first molecular assessment of this group and recovered *Uromyias* nested within *Anairetes*, albeit with weak support. Since Roy et al. (1999), *Uromyias* has been recognized as part of *Anairetes*.

The results of Roy et al. (1999) were empirically inconclusive in placing *Uromyias* with respect to the historically recognized members of *Anairetes*, yet they provided the impetus for dissolving the genus *Uromyias*. In our study, *Uromyias* was independently recovered outside the core *Anairetes* clade by all methods, supporting the validity of two genera and the revival of *Uromyias*. The mtDNA gene tree, concatenated nuDNA tree, and \*BEAST species tree strongly support *Uromyias* sister to *Anairetes*, but our BEST species tree lacks basal resolution and nests *Uromyias* among outgroups. Thus, all data sets and analyses point to the placement of *Uromyias* as basal to the core *Anairetes* clade or outside of it. I therefore advocate the resurrection of the genus *Uromyias* as distinct from *Anairetes*.

I caution that our mtDNA analysis was the only one that included both species of *Uromyias*, and that incomplete lineage sorting in the mitochondrial genome can result in the mtDNA topology being inconsistent with the true species tree. Accordingly, a sizeable mtDNA alignment could lead to high confidence in an “anomalous” gene tree. However, the absence of nuclear sequences for *A. agilis* is not likely to be problematic for three reasons: (1) all evidence from morphology, habitat, distributions, and vocalizations supports the monophyly of *A. agraphia* and *A. agilis*; (2) the branch that subtends *A. agraphia* and *A. agilis* in the mtDNA topology is long, suggesting that incomplete lineage sorting is unlikely; and (3) the high concordance between our mtDNA gene tree and our \*BEAST species tree suggests that the mitochondrial genome is tracking the true species tree.

The discordance among our nuclear gene trees indicates incomplete lineage sorting in nuclear loci and rapid diversification of *Anairetes* and its allies, at least relative to the rate of nuDNA sorting (Degnan and Rosenberg, 2009). Given the evidence to date (phenotypic and genotypic) and convergence problems in BEST, the single most credible hypothesis is that

*Uromyias* and *Anairetes* are monophyletic and that the mitochondrial tree and \*BEAST species tree best reflects interspecific relationships. The mitochondrial genome has an effective population size one-fourth that of autosomal nuclear loci, resulting in a higher probability of tracking the species tree because of a shorter sorting time (Moore, 1995), and there were no apparent convergence issues in our \*BEAST analyses. I have no evidence to suspect that these results are affected by potential problems associated with mitochondrial data such as introgression or saturation/homoplasy. The similar level of resolution between MP and BI analyses from our mtDNA dataset suggests limited homoplasy, and the majority of nodes within our mtDNA topology are corroborated by independent nuclear data and the two species-tree analyses.

It is worth noting the emergent signal in our nuclear concatenated topology. Analyses of the concatenated nuDNA not only recovered bipartitions with increased nodal support, but recovered one bipartition not strongly supported by any individual gene tree. This bipartition was in direct conflict with the IRF2 and PER2 gene trees. IRF2 recovered *A. reguloides* as monophyletic with a posterior probability of 0.97, and PER2 recovered *A. reguloides reguloides* sister to *A. nigrocristatus* with posterior probability of 0.92. The concatenated nuDNA topology recovered the other possible resolution of this triad: *A. nigrocristatus* sister to *A. reguloides albiventris* with posterior probability of 0.95. It appears that the individual nuclear genes do not contain sufficient variation to statistically resolve certain bipartitions, but phylogenetic signal becomes evident when they are combined (Barrett et al. 1991, Chippindale and Wiens 1994). The observed pattern suggest the two bifurcation events in this triad clade occurred in rapid succession and were accompanied by incomplete lineage sorting at most nuclear loci.

#### 4.2.2 Biogeography

The placement of *Uromyias* outside *Anairetes* implies an ecological dichotomy between humid and dry clades with the dry-tolerant southern *Anairetes* clade more prone to morphological, ecological, and lineage diversification. Members of *Uromyias* are habitat specialists that are restricted to stands of humid, *Chusquea* dominated cloud-forest on the east slope of the Andes (1800-3600m elevation; Parker and O'Neill, 1980; del Hoyo et al., 2004; Schulenberg et al., 2007). The *Uromyias* clade, containing only two species and four subspecies (Fig. 5; Dickinson, 2003), has undergone little net ecological and lineage diversification compared to *Anairetes sensu stricto*. *A. agraphia* is endemic to Peru and is replaced north of the Marañon Gap by *A. agilis*, whose distribution extends north to Venezuela. Taxa within *Uromyias* exhibit little difference in morphology and habitat preference. At similar montane elevations, members of *Anairetes sensu stricto* occupy seasonally arid rain-shadow valleys and dry west-facing slopes. Since the most recent common ancestor between the two clades, net diversification within *Anairetes* has yielded six species and 13 subspecies, more than three times the diversity of *Uromyias* (Fig. 4). The difference in clade size when considering subspecies taxa is significant by a binomial test ( $p=0.02$ ); however, I recognize that this level of clade asymmetry is highly likely to occur by chance (Slowinski and Guyer 1993), and statistically robust comparison of diversification rates between humid-restricted and dry-tolerant lineages will require consideration of numerous co-distributed clades.

*Anairetes* has undergone more ecological divergence (i.e. niche divergence) than *Uromyias*, as evident by up to three species occurring in sympatry and *A. parulus* and *A. flavirostris* having become migratory in the southern parts of their range. There is one island (strictly lowland) species, *A. ferndandezianus*, and the basal positions of strictly upper montane *A. nigrocristatus* and *A. alpinus* suggest that *A. reguloides*, *A. parulus*, and *A. flavirostris* have

secondarily invaded the lowlands only on the extreme dry west slope of the Andes and at southern latitudes. The observed difference in net diversification between *Anairetes* and *Uryomias* may be associated with the dry-tolerant vs. humid-restricted dichotomy, but investigation of additional clades that contain dry-tolerant and humid-restricted taxa will be needed to understand whether these disparate ecological pressures generally promote differences in diversification rate in Andean birds. The same ecological dichotomy, reinforced by apparent phylogenetic inertia in humidity-tolerance, is evident in published phylogenies for comparable Neotropical flycatcher clades, including *Elaenia* (Rheindt et al. 2008) and *Muscisaxicola* (Chesser, 2000).

There are two major reasons to predict that ecological generalist lineages will have undergone increased net diversification. First, generalists are more likely to persist through climate shifts and vicariant events due to their broader distributions, larger population sizes (Newman and Pilson, 1997), broader physiological limits (Kellerman et al., 2009), and lower demographic variability (Maliakal-Witt 2004). Second, taxa on the west slope of the Andes and in rain-shadow valleys generally have broader elevational distributions, and these environments exhibit greater climatic heterogeneity across latitudes, elevations, and seasons (Sarmiento, 1986). Dry-tolerant species such as *Anairetes* that span these large gradients in temperature and elevation may be prone to disruptive selection because of disparate ecological pressures, facilitating diversification during periods of isolation (Sargent and Otto, 2006).

## **5. Conclusions**

In comparing our study with Roy et al. (1999) I can ask what phylogenetic advancement over the past decade has had the most impact on inferring species relationships. The major improvement in resolution was garnered by sequencing nearly fourfold more basepairs of

mtDNA. Resolution dramatically increased (from two to six nodes strongly supported at the interspecific level within the *Anairetes* complex), while using the same phylogenetic method (MP) and locus (mtDNA). Each subsequent advancement (partitioned model-based Bayesian analyses, inclusion of multiple unlinked nuclear loci, and species-tree analyses) recovered topologies that were highly consistent with the mtDNA tree under MP (1, 0, and 0 additional resolved nodes, respectively). Importantly, the latter developments provide more robust and easily interpretable empirical support for species relationships, providing an improved platform for taxonomic revision and biogeographic interpretation.

The ongoing shift from the multi-locus concatenation paradigm to a gene-tree coalescent approach allows phylogeneticists to account for conflicting signals due to ancestral polymorphism. BEST and \*BEAST recovered a topology that was highly congruent with other methods and should provide a more realistic picture of nodal support and branch lengths (Edwards, 2009). However, our study and other studies report convergence problems between independent runs in BEST. Convergence problems appeared to affect basal topology more than the core *Anairetes* topology, which was strongly corroborated by both runs.

The mtDNA and \*BEAST species-tree topologies appear to provide a robust and credible hypothesis for the relationships among *Anairetes* taxa (Figure 4a, Figure 3b). This leading hypothesis is not contradicted by any strongly supported nodes for individual nuclear loci and it corroborates the theory that haploid mtDNA is more likely to have correctly tracked the species tree across rapid successive speciation events. *A. agraphia* and *A. agilis* were consistently recovered outside *Anairetes sensu stricto*, supporting the validity of the genus *Uromyias*. The sister relationship of *Uromyias* to *Anairetes* suggests an ecological dichotomy between humid-restricted and dry-tolerant clades, with the latter having undergone more morphological,

ecological, and lineage diversification. The general importance of this ecological dichotomy to diversification is a key biogeographical question to be addressed as phylogenies of additional Andean bird groups emerge.

### **Acknowledgements**

This study would not have been possible without vouchered tissue samples provided by the Academy of Natural Sciences, Philadelphia (Nate Rice), LSU Museum of Natural Sciences (Robb Brumfield and Donna Dittman), and the American Museum of Natural History (Joel Cracraft, Peter Capainolo, and Thomas Trombone). This work was funded in part by NSF DEB-1146491, NCCR-NIH P20RR018754, the UNM Center for Evolutionary and Theoretical Immunology (CETI), The American Ornithologists' Union, and the Frank M. Chapman Memorial Fund of the American Museum of Natural History. I thank Phred M. Benham, Natalie A. Wright, Matthew R. Jones, Emil Bautista Obispo, Homan Castillo, Andrew B. Johnson, C. Jonathan Schmitt, Donna C. Schmitt, Glenn F. Seeholzer, and Centro de Ornitología y Biodiversidad (CORBIDI; Thomas Valqui, Emil Bautista Obispo, Willy Nanez) for fieldwork and logistical support to obtain key samples for this study. I thank Liang Liu, Adam Leache and Matt Fujita for advice in implementing BEST. Comments from Steve Poe, Elizabeth J. Beckman, Phred M. Benham, Natalie A. Wright, Sabrina M. McNew, Jose G. Tello, and two anonymous reviewers greatly improved this manuscript.

### **Literature cited**

Akaike, H., 1974. A new look at the statistical model identification. *IEEE Trans Auto Control*. 19, 716-723.

- Alstrom, P., Fregin, S., Norman, J. A., Ericson, P. G. P., Christidis, L., Olsson, U., 2011. Multilocus analysis of a taxonomically densely sampled dataset reveal extensive non-monophyly in the avian family Locustellidae. *Mol. Phylogent. Evol.* 58, 513-526.
- Baker, R. H., Yu, X. B., DeSalle, R., 1998. Assessing the relative contribution of molecular and morphological characters in simultaneous analysis trees. *Mol. Phylogent. Evol.* 9, 427-436.
- Barrett, M., Donoghue, M. J., Sober, E., 1991. Against consensus. *Syst. Zool.* 40, 486-493.
- Beiko, R. G., Keith, J. M., Harlow, T. J., Ragan, M. A., 2006. Searching for convergence in phylogenetic Markov chain Monte Carlo. *Syst. Biol.* 55, 553-565.
- Belfiore, N. M., Liu, L., Moritz, C., 2008. Multilocus phylogenetics of a rapid radiation in the genus *thomomys* (Rodentia : Geomyidae). *Syst. Biol.* 57, 294-310.
- Bonier, F., Martin, P. R., Moore, I. T., 2008. First description of the nest and young of the Agile Tit-Tyrant (*Uromyias agilis*). *Ornitologia Neotropical* 19, 117-122.
- Boyle, W. A., Conway, C. J., 2007. Why migrate? A test of the evolutionary precursor hypothesis. *Am. Nat.* 169, 344-359.
- Brandley, M. C., Schmitz, A., Reeder, T. W., 2005. Partitioned Bayesian analyses, partition choice, and the phylogenetic relationships of scincid lizards. *Syst. Biol.* 54, 373-390.
- Braun, G., 1997. The use of digital methods in assessing forest patterns in an Andean environment: The *Polylepis* example. *Mountain Research and Development* 17, 253-262.
- Brooke, M. L., 1987. The birds of the Juan Fernandez Islands, Chile. ICBP Study Report, 1-50.
- Brownstone, D., Valletta, R., 2001. The bootstrain and multiple imputations: Harnessing increased computing power for improved statistical tests. *J. Econ. Perspect.* 15, 129-141.

- Brumfield, R. T., Liu, L., Lum, D. E., Edwards, S. V., 2008. Comparison of Species Tree Methods for Reconstructing the Phylogeny of Bearded Manakins (Aves: Pipridae, Manacus) from Multilocus Sequence Data. *Syst. Biol.* 57, 719-731.
- Bull, J. J., Huelsenbeck, J. P., Cunningham, C. W., Swofford, D. L., Waddell, P. J., 1993. Partitioning and combining data in phylogenetic analysis. *Syst. Biol.* 42, 384-397.
- Carriker, M. A., 1933. Descriptions of new birds from Peru, with notes on other little-known species. *Proc. Natl. Acad. Sci. USA* 85, 1-38.
- Chapman, F. M., 1919. Descriptions of proposed new birds from Peru, Bolivia, Argentina, and Chile. *B. Am. Mus. Nat. Hist.* 41, 323-333.
- Chesser, R. T. 1999. Molecular systematics of the rhinocryptid genus *Pteroptochos*. *Condor* 101, 439-446.
- Chesser, R.T., 2000. Evolution in the High Andes: the Phylogenetics of Muscisaxicola Ground-Tyrants. *Mol. Phylogent. Evol.* 15, 369-380.
- Chippindale, P. T., Wiens, J. J., 1994. Weighting, partitioning, and combining characters in phylogenetic analysis. *Syst. Biol.* 43, 278-287.
- Clements, J., 2004. *The Clements Checklist of the Birds of the World* (5 ed.). Ithaca, New York, Cornell University Press.
- Cornelius, C., Cofre, H., Marquet, P. A., 2000. Effects of habitat fragmentation on bird species in a relict temperate forest in semiarid Chile. *Conserv. Biol.* 14, 534-543.
- Cranston, K. A., Hurwitz, B., Ware, D., Stein, L., Wing, R. A. (2009). Species Trees from Highly Incongruent Gene Trees in Rice. *Syst. Biol.* 58, 489-500.
- Degnan, J. H., Rosenberg, N. A., 2009. Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends Ecol. Evol.* 24, 332-340.

- Degnan, J. H., Salter, L. A., 2005. Gene tree distributions under the coalescent process. *Evolution* 59, 24-37.
- Del Hoyo, J., Elliot, A., Sargatal, J., Christie, D. A. 2004. Handbook of the birds of the world, Vol. 9. Cotingas to Pipits and Wagtails. Lynx, Barcelona.
- Dickinson, E. C., 2003. The Howard and Moore Complete Checklist of the Birds of the World. Christopher Helm, London.
- Dingle, C., Lovette, I. J., Canaday, C., Smith, T. B., 2006. Elevational zonation and the phylogenetic relationships of the Henicorhina wood-wrens. *Auk* 123, 119-134.
- Drummond, A. J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* 7, 214.
- Drummond, A. J., Ho, S. Y. W., Rawlence, N., Rambaut, A., 2007. A Rough Guide to BEAST 1.4. [http://beast-mcmc.googlecode.com/files/BEAST14\\_Manual\\_6July2007.pdf](http://beast-mcmc.googlecode.com/files/BEAST14_Manual_6July2007.pdf).
- Edgar, R. C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792-1797.
- Edwards, S. V., 2009. Is a new and general theory of molecular systematic emerging? *Evolution* 63, 1-19.
- Felsenstein, J., 1985. Confidence-limits on phylogenies: An approach using the bootstrap. *Evolution* 39, 783-791.
- Fjeldså, J., Krabbe, N., 1990. *Birds of the High Andes*, Apollo Books, Denmark.
- Gering, E.J., Opazo, J.C., Storz, J.F., 2009. Molecular evolution of cytochrome b in high- and low-altitude deer mice (genus *Peromyscus*). *Heredity* 102, 226-235.
- Gill, F., Donsker, D., 2008. IOC World Bird Names (version 1.5).  
<<http://www.worldbirdnames.org>>

- Hall, B. G., 2007. Phylogenetic trees made easy: a how-to manual 3<sup>rd</sup> ed. Sinauer Associates, Sunderland, MA.
- Harshman, J., 1996. Phylogeny, evolutionary rates, and ducks. Unpublished Ph.D. Dissertation, University of Chicago.
- Heled, J., Drummond, A. J., 2010. Bayesian inference of species trees from multilocus data. *Mol. Biol. Evol.* 27, 570-580.
- Hellmayr, C. E., 1927. Catalogue of birds of the Americas and the adjacent islands in Field Museum of Natural History. *Field Mus. Nat. Hist. Publ.* 242 Zool. Ser. 13, 1-517.
- Hillis, D. M., 1987. Molecular versus morphological approaches to systematic. *Annu. Rev. Ecol. Syst.* 18, 23-42.
- Huelsenbeck, J. P., Crandall, K. A., 1997. Phylogeny estimation and hypothesis testing using maximum likelihood. *Annu. Rev. Ecol. Syst.* 28, 437-466.
- Huelsenbeck, J. P., Rannala, B., 1997. Phylogenetic methods come of age: Testing hypotheses in an evolutionary context. *Science* 276, 227-232.
- Huelsenbeck, J. P., Ronquist, F., 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17, 754-755.
- Jaramillo, A., 2003. *Birds of Chile*. Princeton University Press, Princeton, NJ.
- Johnson, A. B., McNew, S. M., Graus, M. S., Witt, C. C., 2011. Mitochondrial DNA and meteorological data suggest a Caribbean origin for New Mexico's first Sooty Tern (*Onychoprion fuscatus*). *Western Birds* 42, 233-242.
- Kimball, R. T., Braun, E. L., Barker, F. K., Bowie, R., Braun, M. J., Chojnowski, J. L., Hackett, S. J., Han, K. L., Harshman, J., Heimer-Torres, V., Holznagel, W., Huddleston, C. J., Marks, B. D., Miglia, K. J., Moore, W. S., Reddy, S., Sheldon, F. H., Smith, J. V., Witt,

- C. C., Yuri, T. 2008. A well-tested set of primers to amplify regions spread across the avian genome. *Mol. Phylogent. Evol.* 50, 654-660.
- Kellerman, V., B. van Heerwaarden, C. M. Sgró, A. A. Hoffmann. 2009. Fundamental evolutionary limits in ecological traits drive *Drosophila* species distributions. *Science* 325:1244-1246.
- Kocher, T. D., Thomas, W. K., Meyer, A., Edwards, S. V., Paabo, S., Villablanca, F. X., Wilson, A. C., 1989. Dynamics of mitochondrial-NDA evolution in animals: amplification and sequencing with conserved primers. *Proc. Natl. Acad. Sci. USA* 86, 6196-6200.
- Kolaczkowski, B., Thornton, J. W., 2004. Performance of maximum parsimony and likelihood phylogenetics when evolution is heterogeneous. *Nature* 431, 980-984.
- Lanyon, N. E., 1988. A phylogeny of the thirty-two genera in the Elaenia assemblage of tyrant flycatchers. *Am. Mus. Novit.* 1-57.
- Leache, A. D., Fujita, M. K., 2010. Bayesian species delimitation in West African forest geckos (*Hemidactylus fasciatus*). *P. Roy. Soc. B-Biol. Sci.* 277, 3071-3077.
- Linnen, C. R., Farrell, B. D., 2008. Comparison of Methods for Species-Tree Inference in the Sawfly Genus *Neodiprion* (Hymenoptera: Diprionidae). *Syst. Biol.* 57, 876-890.
- Liu, L., 2008. BEST: Bayesian estimation of species trees under the coalescent model. *Bioinformatics* 24, 2542-2543.
- Liu, L., Pearl, D. K., 2007. Species trees from gene trees: Reconstructing Bayesian posterior distributions of a species phylogeny using estimated gene tree distributions. *Syst. Biol.* 56, 504-514.
- Liu, L., Pearl, D. K., Brumfield, R. T., Edwards, S. V., 2008. Estimating species trees using multiple-allele DNA sequence data. *Evolution* 62, 2080-2091.

- Lloyd, H., 2008. Influence of within-patch habitat quality on high-Andean *Polylepis* bird abundance. *Ibis* 150, 735-745.
- Maddison, D. R., Maddison, W. P., 2005. *MacClade*. Sinauer Associates, Sunderland, MA.
- Maliakal-Witt, S. 2004. Microhabitat Distribution and Demography of Two Florida Scrub Endemic Plants with Comparisons to Their Habitat-Generalist Congeners. Ph.D. dissertation, Louisiana State Univ., Baton Rouge, LA.
- Matsen, F. A., Steel, M., 2007. Phylogenetic mixtures on a single tree can mimic a tree of another topology. *Syst. Biol.* 56, 767-775.
- Mindell, D. P., Sorenson, M. D., Dimcheff, D. E., 1998. An extra nucleotide is not translated in mitochondrial ND3 of some birds and turtles. *Mol. Biol. Evol.* 15, 1568-1571.
- Moore, W. S., 1995. Inferring phylogenies from mtDNA variation: Mitochondrial-gene trees versus nuclear-gene. *Evolution* 49, 718-726.
- Moritz, C., Patton, J. L., Schneider, C. J., Smith, T. B., 2000. Diversification of rainforest faunas: An integrated molecular approach. *Annu. Rev. Ecol. Syst.* 31, 533-563.
- Newman, D., Pilson, D., 1997. Increase probability of extinction due to decreased genetic effective population size: Experimental populations of *Clarkia Pulchella*. *Evolution* 51, 354-362.
- Nylander, J. A. A., 2004. *MrModeltest v2*. Program Distributed by the Author. Evolutionary Biology Centre, Uppsala University.
- Ohlson, J., Fjeldsa, J., Ericson, P. G. P., 2008. Tyrant flycatchers coming out in the open: phylogeny and ecological radiation of Tyrannidae (Aves, Passeriformes). *Zool. Scr.* 37, 315-335.

- Parker, T. A., III, O'Neill, J. P., 1980. Notes on little known birds of the upper Urubamba Valley, southern Peru. *Auk* 97, 167-176.
- Patterson, S. A., Morris-Pocock, J. A., Friesen, V. L., 2011. A multilocus phylogeny of Sulidae (Aves: Pelecaniformes). *Mol. Phylogenet. Evol.* 58, 181-191.
- Rada, F., Azocar, A., Briceno, B., Gonzalez, J., GarciaNunez, C., 1996. Carbon and water balance in *Polylepis sericea*, a tropical treeline species. *Trees-Structure and Function* 10, 218-222.
- Rambaut, A., Drummond, A. J., 2009. Tracer version 1.5.0 <<http://beast.bio.ed.ac.uk>>.
- Remsen, J. V., Jr., Cadena, C. D., Jaramillo, A., Nores, M., Pacheco, J. F., Pérez-Emán, J., Robbins, M. B., Stiles, F. G., Stotz, D. F., Zimmer, K. J., 2003. A classification of the bird species of South America. American Ornithologists' Union. <<http://www.museum.lsu.edu/~Remsen/SACCBaseline.html>>
- Rheindt, F. E., Christidis, L., Norman, J. A., 2008. Habitat shifts in the evolutionary history of a Neotropical flycatcher lineage from forest and open landscapes. *BMC Evolutionary Biology* 8, 193.
- Ronquist, F., Huelsenbeck, J. P., 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19, 1572-1574.
- Roy, M. S., Torres-Mura, J. C., Hertel, F., 1999. Molecular phylogeny and evolutionary history of the tit-tyrants (Aves : Tyrannidae). *Mol. Phylogent. Evol.* 11, 67-76.
- Rozen, S., Skaletsky, H., 2000. Primer3 on the WWW for general users and for biologist programmers, in: Krawetz, S., Misener, S. (Eds.), *Bioinformatics Methods and Protocols* in the series *Methods in Molecular Biology*. Humana Press, Totowa, NJ, pp. 365-386.

- Sargent, R. D., Otto, S. P., 2006. The role of local species abundance in the evolution of pollinator attraction in flowering plants. *Am. Nat.* 167, 67-80.
- Sarmiento, G., 1986. Ecologically crucial features of climate in high tropical mountains, in: Vuilleumier, F., Monasterio, M. (Eds.), *High Altitude Tropical Biogeography*, Oxford University Press, Oxford, pp. 11-45.
- Schulenberg, T. S., Stotz, D. F., Lane, D. F., O'Neill, J. P., Parker, T. A., III. 2007. *Birds of Peru*. Princeton University Press, New Jersey.
- Sclater, P.L., 1888. Catalogue of the birds in the British Museum. *Br. Mus. Nat. Hist.* 14, 1-494.
- Sehgal, R. N. M., Lovette, I. J., 2003. Molecular evolution of three avian neurotrophin genes: Implications for proregion functional constraints. *J. Mol. Evol.* 57, 335-342.
- Slowinski, J. B., Guyer, C., 1993. Testing whether certain traits have caused amplified diversification: An improved method based on a model of random speciation and extinction. *Am. Nat.* 142, 1019-1024.
- Smith, W. J., 1971. Behavioral characteristics of Serpophaginine tyrannids. *Condor* 73, 259-&.
- Sorenson, M. D., Ast, J. C., Dimcheff, D. E., Yuri, T., Mindell, D. P., 1999. Primers for a PCR-based approach to mitochondrial genome sequencing in birds and other vertebrates. *Mol. Phylogent. Evol.* 12, 105-114.
- Steel, M., Penny, D., 2000. Parsimony, likelihood, and the role of models in molecular phylogenetics. *Mol. Biol. Evol.* 17, 839-850.
- Stotz, D. F., Fitzpatrick, J. W., Parker, T. A., III, Moskovits, D. K., 1996. *Neotropical Birds: Ecology and Conservation*. University of Chicago Press, Chicago, IL.
- Swofford, D. L., 2002. *PAUP\*: phylogenetic analysis using parsimony (\*and other methods) version 4.0b10*. Sinauer Associates, Sunderland, MA.

- Tello, J. G., Moyle, R. G., Marchese, D. J., Cracraft, J., 2009. Phylogeny and phylogenetic classification of the tyrant flycatchers, cotingas, manakins, and their allies (Aves: Tyrannides). *Cladistics* 25, 429-467.
- Thomson, R. C., Shedlock, A. M., Edwards, S. V., Shaffer, H. B., 2008. Developing markers for multilocus phylogenetics in non-model organisms: A test case with turtles. *Mol. Phylogent. Evol.* 49, 514-525.
- Traylor, M. A. J., 1977. A classification of the tyrant flycatchers Tyrannidae. *Bull. Mus. Compara. Zool.* 148, 129-184.
- Waters, J. M., Rowe, D. L., Burrige, C. P., Wallis, G. P., 2010. Gene Trees versus Species Trees: Reassessing Life-History Evolution in a Freshwater Fish Radiation. *Syst. Biol.* 59, 504-517.
- Webb, T. J., Gaston, K. J., 2003. On the heritability of geographic range sizes. *Am. Nat.* 161, 553-566.

**Table 1.** Tissue samples included in this study. Specimen-vouchered museum archived tissues were used when possible. Museums include ANSP: Academy of Natural Sciences, Philadelphia, Pennsylvania, USA; AMNH: American Museum of Natural History, New York, New York, USA; LSUMNS: Louisiana State University Museum of Natural Science, Baton Rouge, Louisiana, USA; MNHN: Museo Nacional de Historia Natural, Santiago, Chile; MSB: Museum of Southwestern Biology, University of New Mexico, Albuquerque, New Mexico, USA; NMR: Swedish Museum of Natural History, Stockholm, Sweden; ZMUC: Zoological Museum, University of Copenhagen, Copenhagen, Denmark. I used Genbank sequences for *A. agilis*, *A. fernandezianus* (Roy et al, 1999), and *S. munda* (Ohlson et al., 2008).

Taxon	Museum & Catalog No.	Locality	Accession number
<i>Anairetes agilis</i>	ANSP-18890	Loja, Ecuador	Pending-pending
<i>A. agilis</i>	ZMUC	Loja, Ecuador	AF066998
<i>A. agraphia agraphia</i>	MSB:Bird:34165	Cusco, Peru	Pending-pending
<i>A. alpinus bolivianus</i>	MSB:Bird:27098	Cusco, Peru	Pending-pending
<i>A. fernandezianus</i>	MNHN	Robinson Crusoe Island, Chile	AF066992, AF067001
<i>A. flavirostris arequipae</i>	LSUMNS-B103867	Arequipa, Peru	Pending-pending
<i>A. flavirostris flavirostris</i>	AMNH-10403	Anelo, Argentina	Pending-pending
<i>A. flavirostris cuzcoensis</i>	MSB:Bird:33661	Apurimac, Peru	Pending-pending
<i>A. flavirostris huancabambae</i>	MSB:Bird:34056	Piura, Peru	Pending-pending
<i>A. nigrocristatus</i>	MSB:Bird:36059	Ancash, Peru	Pending-pending
<i>A. paralus parulus</i>	AMNH-12192	Chacabuco, Chile	Pending-pending
<i>A. paralus patagonicus</i>	AMNH-10349	Anelo, Argentina	Pending-pending
<i>A. paralus aequatorialis</i>	LSUMNS-B432	Piura, Peru	Pending-pending
<i>A. reguloides reguloides</i>	MSB:Bird:35064	Tacna, Peru	Pending-pending

<i>A. reguloides albiventris</i>	MSB:Bird:32953	Lima, Peru	Pending-pending
<i>Culicivora caudacuta</i>	LSUMNS-B15410	Santa Cruz, Bolivia	Pending-pending
<i>Mecocerculus leucophrys</i>	LSUMNS-B7476	Amazonas, Venezuela	Pending-pending
<i>Polystictus pectoralis</i>	LSUMNS-B38120	Santa Cruz, Bolivia	Pending-pending
<i>Pseudocolopteryx sclateri</i>	LSUMNS-B7617	Beni, Bolivia	Pending-pending
<i>Serpophaga munda</i>	LSUMNS-B37650	Santa Cruz, Bolivia	Pending-pending
<i>Serpophaga munda</i>	NRM-947171	Alto Paraguay, Paraguay	EU231780

---

**Table 2.** Loci included in this study and primers used for PCR amplification and sequencing reactions.

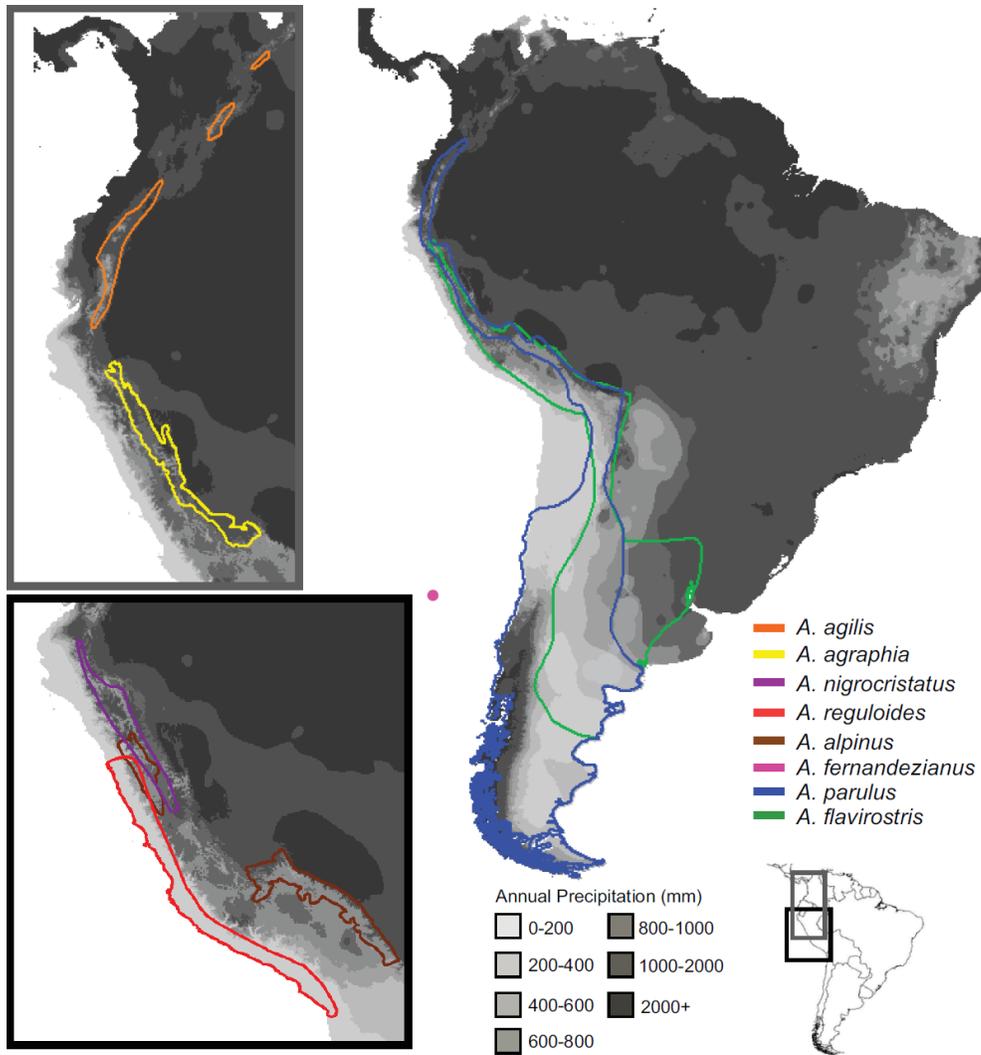
Gene	Length (bp)	Primer name	Sequence (5'-3')	Source
ND2	1021	L5219	CCCATACCCCGAAAATGATG	Sorenson et al. (1999)
		H6313	CTCTTATTTAAGGCTTTGAAGGC	Sorenson et al. (1999)
ND3	351	L10647	TTYGAAGCMGCMGCMTGATSCTG	Mindell et al. (1998)
		H11151	GATTTGTTGAGCCGAAATCAA	Chesser (1999)
Cytb	1000	L14841	GCTTCCATCCAACATCTCAGCATGATGAAA	Kocher et al. (1989)
		H4a	AAGTGGTAAGTCTTCAGTCTTTGGTTTACAAGACC	Harshman (1996)
BDNF	688	ChickBDNF-5'	ATGACCATCCTTTTCCTTACTATG	Sehgal and Lovette (2003)
		ChickBDNF-3'	TCTTCCCCTTTTAATGGTTAATGTAC	Sehgal and Lovette (2003)
IRF2	642	IRF2.2F	ATGTCTTTGGGTCGGGTTTA	Kimball et al. (2008)
		IRF2.3R	GAAACTGGGCAATTCACACA	Kimball et al. (2008)
Myo2	747	E2F1	GAAGATCTGAAGAAACATGGAGCTA	This study
		E3R1	CAATGACCTTGATAATGACTTCAGA	This study
MUSK	641	MUSK- I3F2	AAATAACCCGACCACCTGTAAA	Kimball et al. (2008)

		MUSK- I3R2	TAGGCACTGCCAGACTGTT	Kimball et al. (2008)
		AnaMUSK-F	AAATAATAGAAGGCTTAAAGG	this study
		AnaMUSK-R	CTCTGGACATTGTGTATCCTT	this study
NGF	705	AIINGF5'	GGTGCATAGCGTAATGTCCATG	Sehgal and Lovette (2003)
		AIINGF3'	ATAATTTACAGGCTGAGGTAG	Sehgal and Lovette (2003)
PER2	612	PER.9F	CATCTTCAYCCAAATGACAGACC	Kimball et al. (2008)
		PER.10R	CCTGATTGGTGAATAGTCAAAAGG	Kimball et al. (2008)
		AnaPER2-F	TTTCCGAAAATTAATAAGAATG	this study
		AnaPER2-R	CATACTGAAGAACTGAAATGA	this study

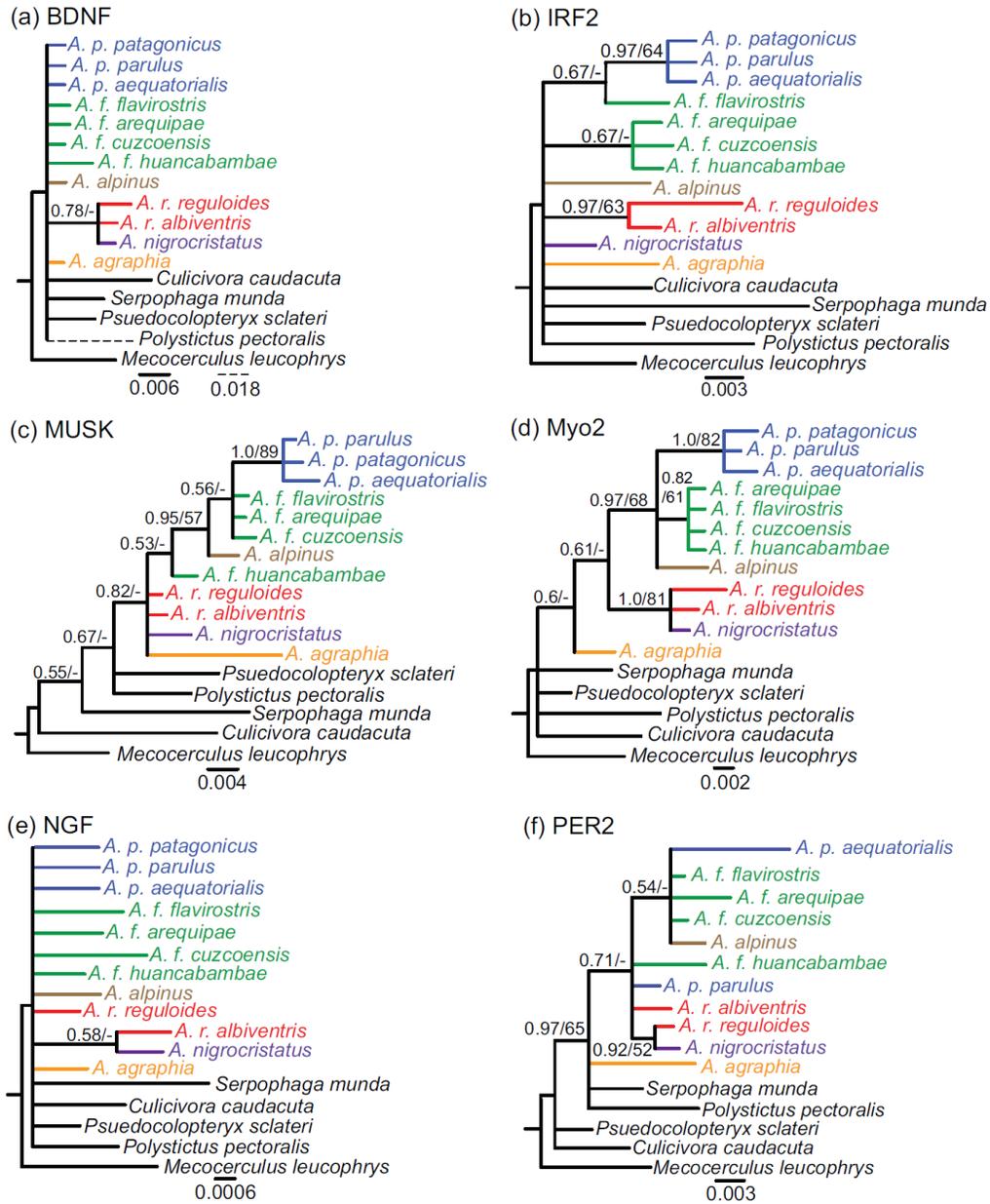
---

**Table 3.** Summary of sequence variation among loci and substitution model recommend by MrModeltest under AIC, implemented in Bayesian gene tree and species tree analyses (MrBayes, BEST, \*BEAST). Partitioning by codon position for mtDNA was only employed in Bayesian gene tree analyses (MrBayes) and \*BEAST specie-tree analyses. Maximum p-distance is among ingroup taxa only (*Anairetes*, including *Uromyias*).

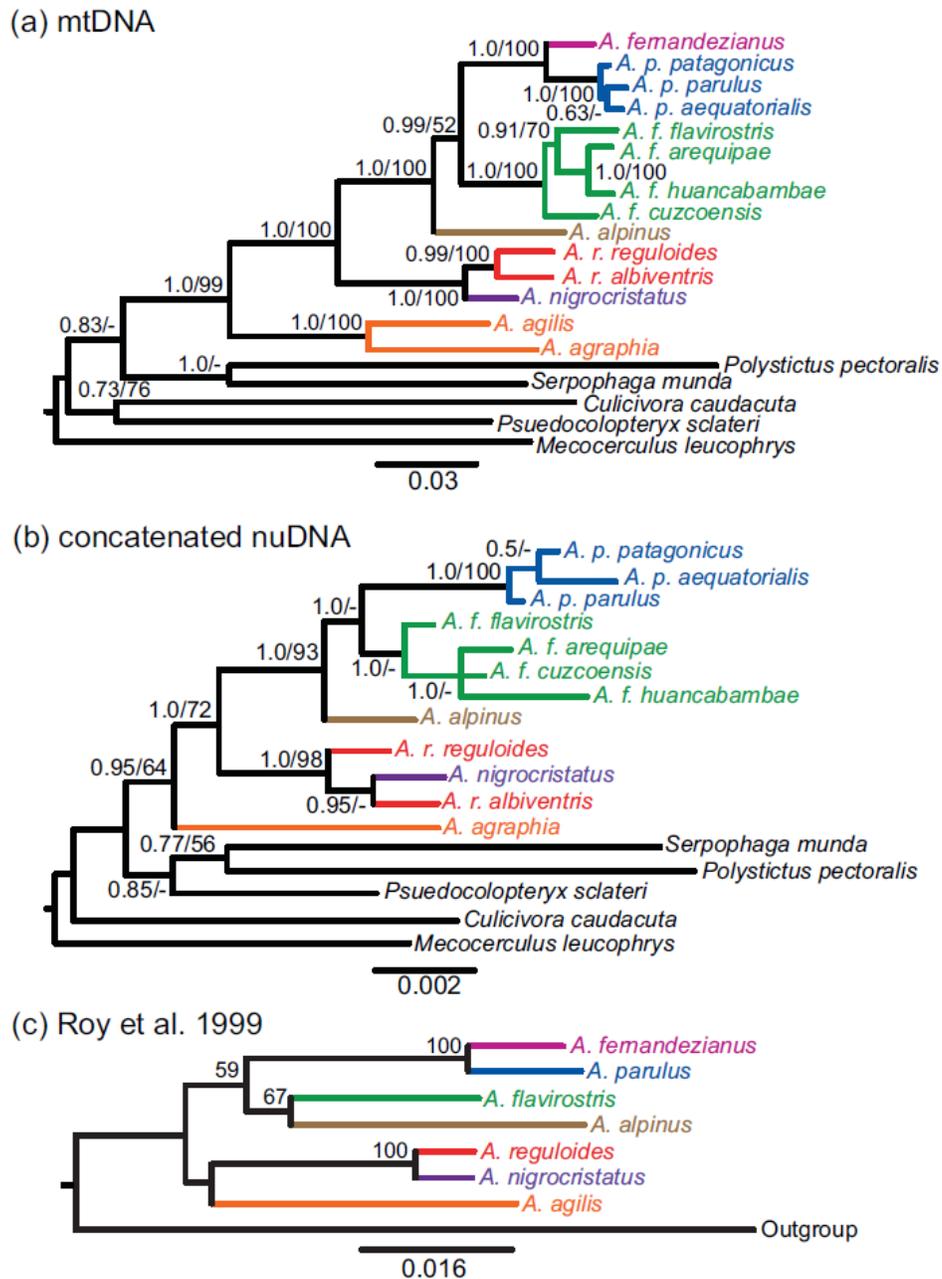
Locus	Number of variable sites	Number of parsimony informative characters	Maximum p-distance	Substitution model	Base frequency (A, C, G, T)
mtDNA	809	539	0.117	-	-
codon pos. 1	164	93	0.091	GTR + I + G	0.2906, 0.2946, 0.1839, 0.2255
codon pos. 2	51	19	0.043	GTR + I	0.1771, 0.3018, 0.1123, 0.4088
codon pos. 3	594	427	0.282	GTR + I + G	0.3879, 0.3679, 0.0384, 0.2057
BDNF	19	7	0.000	HKY + I	0.2755, 0.2632, 0.2745, 0.1867
IRF2	30	10	0.009	HKY + I	0.3373, 0.1888, 0.2056, 0.2684
MUSK	51	10	0.017	HKY + I	0.2902, 0.1809, 0.2225, 0.3064
Myo2	34	12	0.003	HKY + I	0.2907, 0.2190, 0.2402, 0.2501
NGF	8	2	0.000	F81	0.2299, 0.3144, 0.2933, 0.1624
PER2	41	8	0.013	HKY + G	0.3207, 0.1378, 0.1657, 0.3758



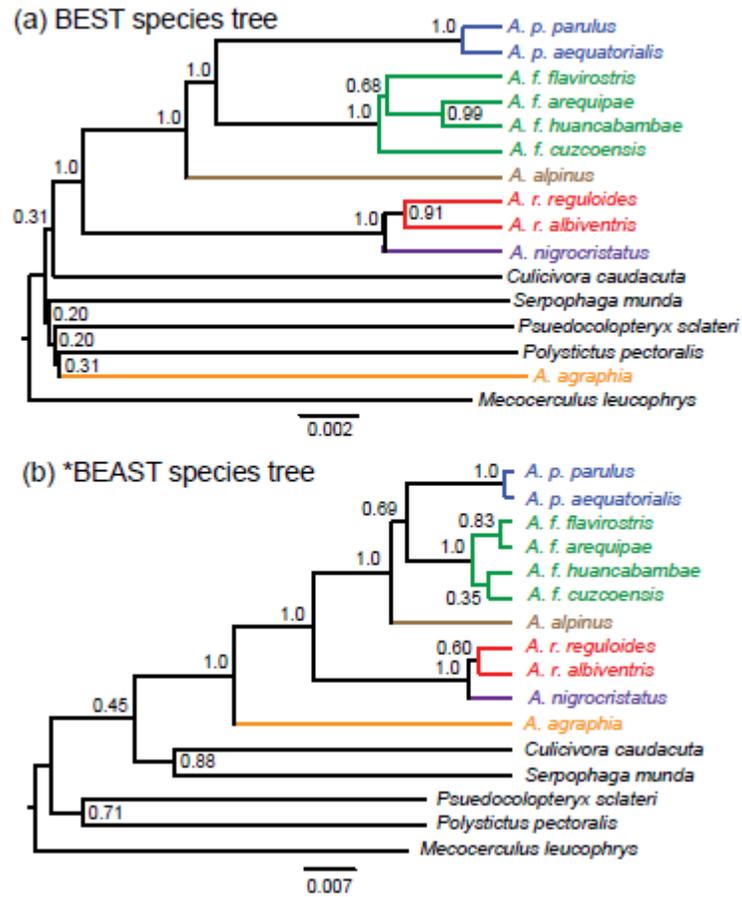
**Figure 1.** Species distributions in South America overlaying mean annual precipitation from ~1950-2000 (Worldclim v.1.3, 2004).



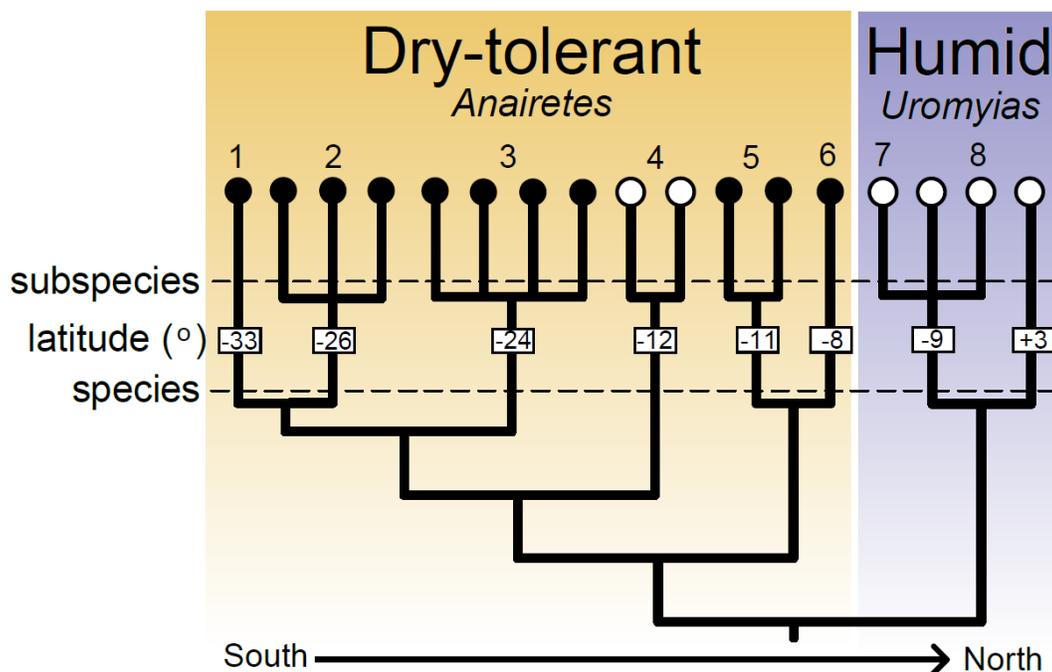
**Figure 2.** Gene tree topologies for each nuclear locus obtained from Bayesian Inference and Maximum Parsimony analyses. Values at nodes refer to posterior probability of BI/MP bootstrap. Nodes were collapsed if posterior probabilities were under 50%.



**Figure 3.** Mitochondrial genome topology and concatenated nuclear topology obtained from Bayesian Inference and Maximum Parsimony analyses. Nodes were collapsed if posterior probabilities were under 50%. Fig. 1a-b, values at nodes refer to posterior probability of BI/MP bootstrap. Fig. 1c is the topology obtained from 632 bps of mtDNA using neighbor joining of Kimura 2-parameter distances from Roy et al. (1999), values at nodes refer to 500 bootstrap replicates. Bars denote to substitutions per site.



**Figure 4.** Species-tree topologies recovered from BEST and \*BEAST. Posterior probabilities are labeled at nodes.



**Figure 5.** Net species and subspecies diversification between the dry-tolerant *Anairetes* clade and humid *Uromyias* clade. Latitude= midpoint of species distribution. Black circles = habitat generalists. Open circles = habitat specialists. Species were categorized as generalists if more than one habitat preference is reported in Stotz et al. (1996), if the species occurs in human-modified environments, and/or if the species utilizes non-native flora. (1) *A. fernandezianus*: endemic to Robinson Crusoe Island, Chile, occurs in native montane evergreen forest, gardens, *Eucalyptus*, and other exotic vegetation (Brooke, 1987). (2) *A. parulus*: has the broadest ecological range spanning dry-torn scrub, elfin forest, disturbed humid scrub, *Polylepis* forest, and temperate forest (Cornelius et al. 2000; Jaramillo, 2003; Schulenberg et al., 2007; Lloyd, 2008). (3) *A. flavirostris*: occurs in desert and thorn scrub to semi-humid coastal Lomas and *Polylepis* forest (Jaramillo, 2003; Schulenberg et al., 2007). (4) *A. alpinus*: *Polylepis* specialist. *Polylepis* habitat is considered semi-humid but experiences water stress during their annual four-month dry season in which relative humidity, atmospheric vapor, and precipitation dramatically

decrease (Rada et al., 1996; Braun, 1997). (5) *A. reguloides*: occurs in dry coastal riparian scrub, cultivated hedgerows, arid thorn scrub, semi-arid montane scrub, and *Polylepis* (Fjeldså and Krabbe, 1990; Schulenberg et al., 2007; Pers. obs.). (6) *A. nigrocristatus*: occurs in composite arid montane scrub of *Lupinus* or *Berberis* along rivers and streams, *Polylepis*, and disturbed scrub around pastoral and agricultural fields (Fjeldså and Krabbe, 1990, Pers. obs.). (7 and 8) *A. agraphia* and *A. agilis*, respectively: *Chusquea* bamboo specialists of humid cloud forest (Fjeldså and Krabbe, 1990, Bonier et al., 2008).

## Chapter 2: Diversification by differential adaptation to altitude in Andean tit-tyrants (Aves, Tyrannidae)

### 1. Introduction

Animal populations that are distributed along steep environmental gradients are subject to opposing evolutionary forces. Local environments drive intrapopulation diversity by favoring alleles that benefit individual fitness under local conditions, irrespective of their fitness consequences at other points along the gradient (Williams, 1966; Kawecki and Elbert, 2004). Gene flow along the axis of the environmental gradient homogenizes variation, thus preventing subpopulations from attaining genotypes that optimize fitness under local prevailing conditions. The outcome of these opposing forces will depend on the strength of diversifying selection and the rate and direction of gene flow along the selection gradient. This selection-migration balance favors local adaptation when strong selection is coupled with low rates of migration (Maynard Smith and Haigh, 1974, Kawecki and Elbert, 2004; Cheviron and Brumfield, 2009). Local adaptation plays a critical role in maintaining intraspecific variation (Felsenstein, 1976; Hedrick et al., 1976; Hedrick, 1986; Schneider et al. 1999) and it may be an important initial step towards speciation (Schluter, 2001; Turelli et al., 2001; Via, 2001). A further possibility is that differential local adaptation promotes diversification at later stages of the speciation processes when incipient species come into contact and interbreed along environmental gradients.

High elevations present metabolic challenges for endothermic organisms that can affect performance and subsequent fitness. Previous studies have shown that abiotic pressures at high elevations can impose incredibly strong selection on parts of the

genome (Cheviron and Brumfield, 2009; McCracken et al., 2009; Storz et al., 2010), making high elevations an ideal system to address the role of local adaptation in speciation. High-elevation pressures include cold temperatures, increased atmospheric radiation, increased desiccation, and most notably hypobaric hypoxia (Monge and Leon-Velarde, 1991; Rezende et al., 2005). Elevations reaching 4000 m exhibit a partial pressure of oxygen ( $PO_2$ ) ~40% less than that of sea level (Beall, 2007). Reduction of  $PO_2$  can hinder oxygen transport to respiring tissues and decrease metabolic output. High elevation organisms would benefit from traits that maximize the efficiency of oxygen uptake, transport, and delivery, providing hypoxia resistance.

Natural selection for hypoxia resistance has been shown in high elevation mammals and birds (McCracken et al., 2009; Storz et al., 2010).  $PO_2$  at a given elevation does not vary temporally and provides a consistent selective pressure on organisms inhabiting a specific elevation. The consistency of a selective force is important in promoting local adaptation as temporal variation opposes local adaptation by favoring generalist phenotypes (Kisdi, 2002). Hemoglobin and the mitochondrion, specifically, have been identified as targets of natural selection in high elevation vertebrates inhabiting cold and hypoxic environments (Snyder, 1981; Chappell et al. 1988; Elinger et al., 2002; Mishmar et al., 2003; Ruiz-Pesini et al. 2004; Storz et al., 2007, 2009; Storz and Kelly, 2008; McCracken et al., 2009). Structural changes in high elevation hemoglobin can increase the uptake and supply of oxygen to respiring tissue under hypoxic conditions (McCracken et al., 2010; Storz et al., 2010). Further, the mitochondrial genome has been shown to be under strong selection in cold and hypoxic environments because of its role

in oxidative phosphorylation and production of adenosine triphosphate (ATP; Ehinger et al., 2002; Mishmar et al., 2003; Ruiz-Pesini et al. 2004; Cheviron and Brumfield, 2009).

I investigated the role of differential adaptation to altitude in maintaining species limits and distribution limits between sister species: the smaller, widespread *Anairetes reguloides* (Pied-crested Tit-tyrant) and the larger, high-elevation *Anairetes nigrocristatus* (Black-crested Tit-tyrant). *A. reguloides* inhabits the dry, west slope of the Peruvian Andes from sea-level to 4200m, and *A. nigrocristatus* inhabits inter-Andean valleys, such as the Marañon, and primarily occurs at upper montane elevations from 3000m-4200m but locally down to 2200m (museum specimen records). *A. reguloides* and *A. nigrocristatus* are arid-tolerant generalists with no apparent differences in habitat preference; they inhabit arid montane scrub, riparian thickets, *Polylepis* woodland, and hedgerows (Stotz et al., 1996; DuBay and Witt, 2012). Diversification within the *A. reguloides/nigrocristatus* complex is compatible with allopatric isolation along the latitudinal axis of the Andean ridge, as is typical of Andean birds (Krabbe and Schulenberg, 1997; DuBay and Witt, 2012). Subsequent range expansion has led to overlap of the two species in central Peru, but *A. reguloides* appears to be restricted to lower elevations in the presence of *A. nigrocristatus* (Schulenberg et al., 2007).

In this study I investigate the role of differential response to altitude in promoting and maintaining elevational replacement between *A. reguloides* and *A. nigrocristatus*. I quantify the phenotypic and genetic structure of *A. reguloides* and *A. nigrocristatus* using likelihood cline analyses, coalescent-based models of gene flow, and assessments of the hematological response to elevation. I predict that *A. nigrocristatus* will exhibit a muted hematological response to ambient hypoxia because its elevational restriction to upper

montane areas may have facilitated local adaptation to low PO<sub>2</sub>. On the other hand, gene flow can inhibit the spread of locally adapted alleles when a species is contiguously distributed across an environmental gradient. Given the >4000 m elevational range of *A. reguloides* I predict that it will exhibit a marked hematological response to hypoxic stress at high elevation because of the homogenizing pressure of gene flow with contiguously distributed lowland conspecifics. Local adaptation to high elevations is unable to develop or persist in *A. reguloides*. Upon secondary contact differential physiological adaptation to altitude would promote competitive displacement of the two species according to elevation, with the caveat that introgressive hybridization has the potential to break down locally coadapted gene complexes, reducing diversity by eventual fusion-extinction.

## **2. Materials and methods**

### **2.1 Respiratory physiology**

I examined the physiological response of *A. reguloides* and *A. nigrocristatus* to low ambient PO<sub>2</sub> at high elevations. I measured hematological parameters associated with blood-oxygen carrying to quantify the changes in blood chemistry across the elevational distribution of each species (*A. reguloides*, n=35; *A. nigrocristatus*, n=11). The parameters I measured were hemoglobin concentration (Hb), hematocrit (Hct), and mean corpuscular hemoglobin concentration (MCHC). Organisms that experience hypoxic stress at high elevations will exhibit elevated levels in these parameters as a compensatory response to low arterial oxygen saturation (Guyton and Richardson, 1961; McGrath and Weil, 1978; Black and Tenney, 1980). I obtained whole blood samples from live birds by venipuncture on the underside of the wing and collection with heparinized microcapillary tubes (Hct) and cuvettes (Hb). Hct (%) was measured after

centrifuging each microcapillary tube for a minimum of five minutes at 13,000 rpm to separate the red blood cells and plasma. Two Hct samples were taken for each bird and the values were averaged. Hb ( $\text{g dl}^{-1}$  blood) was measured on 5ul of blood using a HemoCue Classic hemoglobin photometer. This photometer produces values for avian blood that are approximately  $1.0 \text{ g dl}^{-1}$  greater than those generated using cyanomethaemoglobin spectrophotometry (Simmons and Lill, 2006). All Hb values were thus corrected for avian samples by subtracting  $1.0 \text{ g dl}^{-1}$ . MCHC ( $\text{g dl}^{-1}$  blood) was calculated by dividing Hb by Hct. Birds were bled immediately upon capture as dehydration is known to inflate blood measurements. Only adult birds with 100% skull ossification and no evidence of a bursa were included in physiological analysis. Individuals from localities with intermediate phenotypes excluded.

## 2.2 Contact zone analysis

I discovered a contact zone at intermediate elevations in Ancash, Peru where phenotypically and genetically intermediate birds occur. I sampled an elevational transect across the hybrid zone to investigate the phenotypic turnover and the rate and direction of genetic introgression. I generated a linear transect of 212 km in Google Earth following protocols similar to Porter et al. (2007; Fig. 1). All localities were less than 12 km off the transect line. The transect roughly corresponds to the Rio Santa Valley, Ancash, Peru, the hypothesized dispersal corridor between the two species. Localities 2-6 lie within the Rio Santa Valley. Locality seven is isolated from the Rio Santa Valley by the Cordillera Blanca, the highest range in the Peruvian Andes. Locality one sits ~65km north along the coast from where the Rio Santa empties into the Pacific Ocean. Localities one and seven represent populations of *A. reguloides* and *A. nigrocristatus*, respectively, which lie

outside of the dispersal corridor. The transect also roughly corresponds to an elevational gradient, increasing from locality one to locality seven. Elevations for each locality are: 1= 309m, 2= 357m, 3= 2900-3800m, 4= 3300m, 5= 3500-3700m, 6= 3600m, and 7= 4000m. It is important to note that locality three spans high elevations up a branching valley on the west slope of the Cordillera Negra before the Rio Santa hooks around the northern tip of the range at 800m elevation (Fig. 1). The lowest pass between the east and west sides of the Negra is ~4200m, with a much higher ridgeline average (highpoint >5000m).

### 2.2.1 Phenotypic structure

I examined the phenotypic structure of *A. reguloides* and *A. nigrocristatus* across the linear transect described above. I measured four morphometric character from 61 individuals (average of 8.71 individuals/locality) of *A. reguloides/nigrocristatus*, including mass, wing chord on a closed wing, crest length from the base of the culmen to the tip of the crest, and white tip length of the outer left rectrix along the rachis from where the leading and trailing edges of the feather are white. I included measurements that could potentially respond to environmental gradients, such as mass and wing chord, and measurements that should be the result of genetic history alone (i.e. drift or selection), such as crest length and amount of white in the tail. I also included individuals in our analysis that were sampled away from zones of potential introgression to help estimate the parental range of each morphometric character for the two species.

I estimated cline shape parameters for each morphometric character using ClineFit (Porter et al., 2007). ClineFit uses a numerical maximum-likelihood algorithm to assess the relationship of data over distance. In our phenotypic analyses, the cline center

( $c$ ) is defined by the location along the transect where the given morphometric character most rapidly shifts from one species to the other; the cline width ( $w$ ) is defined by the distance over which this shift occurs (Szymura and Barton, 1986; Cheviron and Brumfield, 2009). The likelihood algorithm implemented in ClineFit assumes a genetic model with binomial variance, and thus I did not calculate log-likelihood confidence intervals for either  $c$  or  $w$  with morphometric data (Cheviron and Brumfield, 2009). To implement ClineFit, individuals were identified as either *A. reguloides* or *A. nigrocristatus* for each morphometric character if their measurement fell within the range of either parental population. The two parental ranges for each character were determined by comparing individuals along the transect to individuals in other parts of their distributions (*A. reguloides* ranges: mass < 8.5g, crest length < 30mm, wing chord < 53mm, white rectrix tip < 6mm; *A. nigrocristatus* ranges: mass > 9g, crest length > 27.5mm, wing chord > 52mm, white rectrix tip > 8.5mm). Crest length and wing chord overlap slightly between species. Individuals with measurements within these zones of overlap were assigned as either *A. reguloides* or *A. nigrocristatus* based on unique plumage motifs (*A. nigrocristatus*= all black crest feathers, *A. reguloides*= black crest feathers with white edging). Individuals with measurements that fell between the parental ranges for mass and white rectrix tip could not be confidently assigned and were excluded from estimation of  $c$  and  $w$ . These intermediate individuals were only present at localities four and five and exclusion is thus unlikely to affect cline estimates for mass and white rectrix tip. Cline analyses were conducted with 300 parameter tries per annealing step, 2000 replicates, and 30 replicates between saves. Only adult birds (100% skull ossification and no evidence of a bursa) from June-August were included in contact

zone analyses and in determining the parental morphometric ranges for *A. reguloides*. Because of a limited sample size of *A. nigrocristatus* from localities away from the contact zone I used measurements of adult birds from different times of the year to determine the parental range of each morphometric character for this species.

## 2.2.2 Genetic structure

### 2.2.2.1 DNA extraction, PCR, and sequencing

Total DNA was extracted from frozen and ethanol preserved skeletal muscle of *A. reguloides* and *A. nigrocristatus* using the DNeasy Tissue Kit (Qiagen, Valencia, CA). Individuals were sampled from the same seven localities along the linear transect described above. I sequenced four loci for each individual, including a mitochondrial gene [982 base pairs cytochrome b (Cyt b)], two autosomal nuclear intron loci [672 bp myoglobin intron2 (Myo2); 487 bp interferon regulatory factor 2 (IRF2)], and a sex-linked nuclear intron locus [498 bp muscle, skeletal, receptor tyrosine kinase (MUSK)]. Primers and conditions for PCR amplification were obtained from the literature as described by DuBay and Witt (2012). I used PCR and sequencing protocols described by Johnson et al. (2011). Unambiguous double-peaks of equal height in nuclear sequences were coded as heterozygous using the IUPAC ambiguity code. Sequences were aligned using MUSCLE 3.7 (Edgar, 2004) and visually inspected in MacClade 4.08 OS X (Maddison and Maddison, 2005).

### 2.2.2.2 Population differentiation and cline analysis

I determined the allelic phase for heterozygous nuclear sequences using the Bayesian probability algorithm of PHASE (Stephens et al., 2001) implemented in DNAsp v.5 (Librado and Rozas, 2009). I used PHASE protocols described by Sequeira et al.

(2011). Individuals for which PHASE could not assign haplotypes with a probability greater than 90% were excluded from further analyses. Haplotype networks were constructed for each of the four loci using the median-joining algorithm implemented in NETWORK 4.6 (Bandelt et al., 1999).

I used the program Structure 2.3.3 (Pritchard et al., 2000) to examine population differentiation of nuclear loci across the contact zone. Structure computes the Bayesian probabilities of species assignment for individuals with admixed ancestry by maximizing Hardy-Weinberg equilibrium and minimizing linkage disequilibrium (Pritchard et al., 2000). Mitochondrial DNA was excluded from Structure analysis because *Anairetes* mtDNA haplogroups are highly divergent compared to nuclear loci, suggesting a dramatically different evolutionary history and rate of evolution (DuBay and Witt, 2012). Structure was run with a two-population admixture model ( $K=2$ ,  $\alpha=1$ ), independent allele frequencies, and no *a priori* population information. Five independent runs were conducted in Structure with 200,000 generations after 100,000 generations of burn-in. Only individuals with data for all three loci were included.

I examined cline shape parameters for each locus using ClineFit (Porter et al., 2007). I estimated cline center ( $c$ ) and width ( $w$ ) to quantify the genotypic shift from *A. reguloides* to *A. nigrocristatus* across the contact zone. I excluded one *Myo2* haplotype from cline analysis because it could not be confidently assigned to either species. This haplotype was present at similar frequencies at the ends of the transect with limited or zero prevalence at intermediate localities, suggesting that sharing of this haplotype is most likely the result of ancestral polymorphism. All other shared haplotypes could be unambiguously assigned to either species. Haplotypes unique to localities four and five,

where intermediate phenotypes occurred, were excluded from cline analyses because their origin could not be confidently identified. Exclusion of these haplotypes should have little effect on cline shape as they are from intermediate localities along the transect and not present at the ends. I report two-unit log-likelihood support limits that are equivalent to a 95% confidence interval for each parameter estimate (Edwards, 1972).

### 2.2.2.3 Gene flow between *A. reguloides* and *A. nigrocristatus*

I assessed introgression patterns and gene flow between *A. reguloides* and *A. nigrocristatus* across the linear transect using a two population Isolation with Migration (IM) model implemented in IMA2 (Hey, 2010). IMA2 employs a Bayesian MCMC method to fit a coalescent model to genotypic data. I estimated effective population size parameters ( $\Theta_r$ ,  $\Theta_n$ ,  $\Theta_A$ ), rates of gene flow ( $m_{r>n}$ ,  $m_{n>r}$ ), and time since divergence ( $t$ ) between populations of *A. reguloides* and *A. nigrocristatus* (see table 4 for specific definitions for parameters). The number of effective immigrants per generation ( $2N_eM$ ) is estimated in IMA2 by multiplying  $\Theta$  by  $m$ . Because IMA2 assumes that loci are selectively neutral I included only nuclear loci in IM analyses; mtDNA in *Anairetes* was highly divergent and previous avian studies suggest that mtDNA may be under selection at high elevations (Cheviron and Brumfield, 2009; McCracken and Wilson, 2011; DuBay and Witt, 2012). Localities 1-3 were designated as *A. reguloides* because every individual had an *A. reguloides* phenotype and mitochondrial genotype. I used the same rationale to designate localities 6-7 as *A. nigrocristatus*. These designations allowed us to assess gene flow along the entire length of the transect (Cheviron and Brumfield, 2009).

Because IMA2 assumes that loci are free from intra-locus recombination I tested for past recombination events using the four-gamete test (Hudson and Kaplan, 1985)

implemented in DNAsp v.5 (Librado and Rozas, 2009). I detected recombination at all nuclear loci. Therefore, loci were pared to the longest non-recombining fragment (Hey, 2007). I used Myo2 positions 65-450, IRF2 positions 1-440, and MUSK positions 1-229. IMa2 requires fully resolved haplotype data and thus only individuals for whom PHASE could assign haplotypes with a probability greater than 90% were included. The final data matrix consisted of 36 individuals in the *A. reguloides* population and eight individuals in the *A. nigrocristatus* population. I defined an infinite-sites substitution model for all loci with a 1.0 inheritance scalar for autosomal loci (Myo2 and IRF2) and a 0.75 inheritance scalar for the Z-linked locus (MUSK). I assumed a mutation rate of  $3.6 \times 10^{-9}$  substitutions/site/year for autosomal loci and  $3.9 \times 10^{-9}$  substitutions/site/year for the Z-linked locus (Axelsson et al., 2004).

IMa2 was initially run with wide priors to obtain upper-bound estimates that encompass the entire posterior distribution of each prior. I set uniform priors for final analyses with upper-bound estimates determined by the preliminary runs ( $\Theta = 3.7$ ,  $t = 3$ ,  $M = 20$ ). I determined the upper-bound for  $t$  on the assumption that splitting time cannot exceed the time to the most recent common ancestor (TMRCA; McCracken and Wilson, 2011). Analysis was run with a burn-in of 200,000 steps and 30 chains with low geometric heating ( $h_a = 0.96$ ,  $h_b = 0.90$ ). Analysis was run until effective sample size (ESS) estimates for each parameter was greater than 100 (Hey, 2005). Autocorrelations, trend plots, and set comparisons were assessed throughout the run, and I conducted multiple runs with identical priors but different random seed numbers to ensure convergence of the parameter estimates. Because replicated runs produced highly similar

parameter estimates I only present results from the longest run (17 million steps, sampling every 100 steps = 170,000 sampled genealogies).

### 3. Results

#### 3.1 Respiratory physiology

Blood-oxygen carrying capacity parameters for *A. reguloides* ranged from 12.1-21.44 g dl<sup>-1</sup> for Hb, 45.5-58.1% for Hct, and 26.52-36.85 g dl<sup>-1</sup> for MCHC across an elevation span of 300-4142m. Blood parameters for *A. nigrocristatus* ranged from 13.9-17.8 g dl<sup>-1</sup> for Hb, 46.5-57.2% for Hct, and 29.49-33.56 g dl<sup>-1</sup> for MCHC across an elevation span of 3400-4000m. The upper limit for all parameters was consistently higher in *A. reguloides* and the lower limits were consistently higher in *A. nigrocristatus*. *A. reguloides* showed an increase in the three parameters that was positively correlated with elevation (Fig. 2). *A. nigrocristatus* showed no apparent relationship between blood parameters and elevation (Fig. 2). *A. nigrocristatus* showed blood parameter levels at high elevations similar to those of *A. reguloides* at low to mid elevations. At extreme elevations (>3800m) *A. reguloides* showed a spike in blood-oxygen carrying capacity parameters that is significantly higher than *A. nigrocristatus* at similar elevations (Hb  $p=0.000136$ , Hct  $p=0.0153$ , MCHC  $p=0.00597$ ).

#### 3.2 Cline zone analysis

##### 3.2.1 Phenotypic structure

I observed a shift in phenotype from *A. reguloides* to *A. nigrocristatus* across the transect (Fig. 3). The centers for the four morphometric clines were all within 13 km from one another, ranging from 107.49 km (wing chord) to 120.25 km (crest length) from locality one (mass cline center= 114.74 km, white rectrix tip cline center= 114.77 km).

The average cline center for the four morphometric characters was 114.31 km from locality one and located between localities four and five. Cline widths ranged from 7.08 km (white rectrix tip) to 44.58 km (wing chord). Mass and crest length had cline widths of 32.24 km and 12.27 km, respectively.

I observed six phenotypically intermediate birds from localities four and five (i.e. birds that definitively identified as *A. reguloides* for one trait and *A. nigrocristatus* for another trait). All measurements are reported in the following order: mass, crest length, wing chord, white rectrix tip. The three intermediate birds from locality four are reported first followed by the three from locality five. MSB:BIRD (NK173908): 8.31g, 26.17mm, 54.15mm, and 7.39mm. MSB:BIRD (NK173914): 7.68g, 22.91mm, 54.36mm, 5.5mm. MSB:BIRD (NK173927): 8.98g, 27.14mm, 55.79mm, 3.63mm. MSB:BIRD (NK 171736): 8.58g, 28.57mm, 51.46mm, 11.08mm. MSB:BIRD (NK 171731): 7.49g, 29.37mm, 54.63mm, 8.98mm. MSB:BIRD (NK 171756): 8.89g, 23.3mm, 55.52mm, 7.33mm. All individuals from localities 1-3 identified as *A. reguloides* for all morphometric characters, and all individuals from localities six and seven identified as *A. nigrocristatus* for all morphometric characters.

### 3.2.2 Genetic structure

I successfully sequenced 83 individuals for Cyt b, 79 individuals for Myo2, 82 individuals for IRF2, and 81 individuals for MUSK across the linear transect. All localities were represented by at least four individuals per locus, averaging 11.6 individuals/locus/locality. All loci were polymorphic with informative SNPs. I observed over 20-fold greater divergence in the mitochondrial DNA than the nuclear DNA. *A. reguloides* and *A. nigrocristatus* mitochondrial haplogroups were 36 base pairs different,

corresponding to ~3.5% mitochondrial divergence. Nuclear loci averaged a difference of 1.7 base pairs (Fig. 4). There were no shared mtDNA haplotypes between morphologically distinct localities [*A. reguloides*, localities 1-3 (n=43); *A. nigrocristatus*, localities 6-7 (n=18)]. 85% of the individuals from locality four had an *A. reguloides* mtDNA haplotype while 15% had an *A. nigrocristatus* haplotype (n=13). Only 33% of individuals from locality five had an *A. reguloides* mtDNA haplotype while 66% had an *A. nigrocristatus* haplotype. (n=9).

Nuclear loci showed limited haplotype sharing between morphologically distinct localities (Myo2=2 shared haplotypes, IRF2=2, MUSK=1; Fig. 4). The nuclear loci showed sporadic structure; the location within the network and the neighbors of a haplotype did not predict the species identity (Fig. 4). I observed two shared haplotypes in Myo2, one of which was found only in localities one, two, three, and seven with frequencies 0.125, 0.167, 0.130, and 0.285, respectively. This haplotype was absent from intermediate localities along the transect. The observed pattern suggests that the sharing of this haplotype is the result of incomplete lineage sorting rather than introgression. This haplotype has most likely been maintained at low frequencies in both species since the time of the ancestral population. The other shared Myo2 haplotype is the dominant *A. reguloides* haplotype found in all but locality seven. Because of geographic proximity I could not rule out introgression as the cause of this shared haplotype. The sporadic nature of the Myo2 haplotype network suggests that this shared haplotype may also be the result of ancestral variance being maintained in both species. The same pattern is true for the shared haplotypes in IRF2 and MUSK. In all further analyses, however, I conservatively

treated these shared haplotypes (excluding the first Myo2 haplotype) as the result of introgression rather than ancestral variance.

Structure analysis of the three nuclear loci assigned all individuals from localities 1-3 as *A. reguloides* with a posterior probability  $>0.95$  (Fig. 5). I consider confident species assignment as a posterior probability  $>0.95$ . Seven of the eight individuals from locality four were confidently assigned as *A. reguloides* (the eighth individual was assigned as *A. reguloides* with  $P=0.931$ ). One individual from locality four was assigned as *A. reguloides* ( $P=0.984$ ) but had an *A. nigrocristatus* mtDNA haplotype. One individual was confidently assigned to each species from locality five while four individuals could not be confidently assigned to either species. Of these ambiguous individuals from locality five, one individual had an *A. reguloides* mtDNA haplotype and the other three had *A. nigrocristatus* mtDNA haplotypes. Eight of the ten individuals from locality six were confidently assigned as *A. nigrocristatus* (two individuals were assigned as *A. nigrocristatus* with  $P=0.927$  and  $P=0.901$ ). All individuals from locality seven were confidently assigned as *A. nigrocristatus*.

I observed dramatic clinal shifts from *A. reguloides* alleles to *A. nigrocristatus* alleles in all loci along the transect (Fig. 6). The center for the mitochondrial haplotype cline was 109.37 km (101.28-120.79 km) from locality one. The cline centers for the three nuclear loci were shifted  $\sim 17$  km east along the transect but within five km from one another [Myo2 cline center= 128.48 km (114.33-146.80 km), IRF2 cline center= 124.00 km (109.51-143.73 km), MUSK cline center= 128.26 km (113.10-146.69 km)]. The average cline center for all loci (122.56 km) was located within the Rio Santa Valley between localities five and six. All loci produced negative lower-bound support values

for cline width, an outcome that describes a case where the cline is flipped (i.e. the left side asymptotes to 1 and the right side asymptotes to 0; personal communication, Adam Porter). A negative width is biologically impossible, but can be produced in ClineFit if the sample size is low enough that a flipped cline cannot be statistically ruled out. In our data, a negative width would assume that *A. reguloides* alleles are dominant on the right side of the cline and that *A. nigrocristatus* alleles are dominant on the left side. All phenotypic, genetic, and distribution data suggest that a flipped cline is biologically incorrect and a negative width can be ruled out. For all loci I report the lower-bound support value for width as 0. The cline widths ranged from 31.59 km (0-58.06 km) for Cyt b to 81.24 km (0-131.72 km) for IRF2. Myo2 and MUSK had cline widths of 58.32 km (0-98.27 km) and 55.07 km (0-92.35 km), respectively.

It is worth noting that coalescent models such as IM scale effective population size parameters ( $\Theta$ ) and gene flow rate parameters ( $m$ ) with mutation rate ( $\Theta = 4N_e\mu$ ,  $m = M/\mu$ ).  $m$  then becomes the ratio of immigration events per mutation.  $m$  in itself is not very useful unless the value is zero in which case gene flow is zero, irrespective of scaling (Hey, 2011). The gene flow rate, however, can be multiplied by the effective population size parameter ( $\Theta$ ) to produce the number of effective genes entering a population per generation through immigration ( $2N_eM$ ), a more biologically informative value. I report highpoint values and 95% highest posterior density (HPD) confidence intervals for all introgression parameter estimates.

The posterior distributions for all parameter estimates showed a single peak rising and falling to zero with the exception of splitting time ( $t$ ) and ancestral population size ( $\Theta_A$ ). Gene flow rates in both directions statistically encompassed zero and thus I could

not confidently distinguish whether the shared variation between *A. reguloides* and *A. nigrocristatus* is the result of introgression or incomplete lineage sorting. The gene flow rate from *A. nigrocristatus* into *A. reguloides* ( $m_{r>n}$ ) was 0.190 (HPD95% = 0.00-9.69). The gene flow rate from *A. reguloides* into *A. nigrocristatus* ( $m_{n>r}$ ) was 1.010 (HPD95% = 0.00-11.43). The estimate of  $\Theta_r$  was 0.2886 (HPD95% = 0.0472-0.8584), and the estimate of  $\Theta_n$  was 0.2169 (HPD95% = 0.0245-1.549). The number of effective immigrants per generation was less than one and statistically overlapped with zero in both directions. The effective number of immigrants per generation of *A. nigrocristatus* into *A. reguloides* ( $2N_eM_{r>n}$ ) was 0.1927 (HPD95% = 0.00-1.019), and the effective number of immigrants per generation of *A. reguloides* into *A. nigrocristatus* ( $2N_eM_{n>r}$ ) was 0.2451 (HPD95% = 0.00-2.659; Fig. 7).

### 3.3 Evidence of hybridization

The presence of intermediate phenotypes at localities four and five is highly suggestive of hybridization between *A. reguloides* and *A. nigrocristatus*. I confirmed hybridization at these localities by discovering individuals that phenotypically identify as one species but genotypically identify as the other. MSB:BIRD (NK173923) from locality four had phenotypic measurements within the range of *A. reguloides* for all characters (mass = 7.71g, crest length = 25.55mm, wing chord = 51.83mm, white rectrix tip = 2.97mm) but had an *A. nigrocristatus* mtDNA haplotype. MSB:BIRD (NK 171731) from locality five had phenotypic measurements within the range of *A. nigrocristatus* for all character excluding mass (mass = 7.49g, crest length = 29.37mm, wing chord = 54.63mm, white rectrix tip = 8.98mm) but had an *A. reguloides* mtDNA haplotype.

## 4. Discussion

#### 4.1 Differential local adaptation

Differential physiological adaptation to local environmental pressures can drive divergence between closely related taxa. Locally adapted taxa have a performance advantage within their respective environments over maladapted competitors with alleles that have originated in external environments (Kawecki and Elbert, 2004). Hematological data suggest that *A. nigrocristatus* is not experiencing hypoxic stress at high elevations and is physiologically specialized on low PO<sub>2</sub>. *A. nigrocristatus* shows similar levels in blood parameters to those of *A. reguloides* under normoxia, providing it with a performance advantage at high elevations. Physiological data of *A. nigrocristatus* are suggestive of structural changes in the hemoglobin protein that make the molecule a more effective oxygen carrier. Studies in other tetrapods have found the alpha and beta globin genes that code for the hemoglobin protein to be under strong selection at high elevations (McCracken et al., 2010; Storz et al., 2010). Single base pair changes within globin coding regions can have structural effects on the active binding site of the hemoglobin protein. These structural changes can increase oxygen loading in one of two ways: (1) oxygen binding affinity increase at the active binding site (McCracken et al., 2010; Storz et al., 2010) or (2) the binding affinity for organic phosphates, such as inositol pentaphosphate (IPP), decreases at the active binding site, resulting in an increase in available sites for oxygen to bind (unpublished data). Organisms with either of these structural changes would be able to effectively bind and deliver more oxygen in hypoxic environments without increasing erythropoiesis (i.e. red blood cell production).

The apparent spike in blood parameters in high elevation *A. reguloides* suggests increased erythropoiesis, a phenotypically plastic response that is indicative of hypoxic

stress (Black and Tenney, 1980). Elevated erythropoiesis effectively increases hemoglobin concentration, hematocrit, and MCHC. *A. reguloides* experiences this plastic response to low ambient  $PO_2$  to compensate for low arterial oxygen saturation. In theory, increased erythropoiesis should allow an organism to increase their arterial oxygen saturation and supply more oxygen to deprived tissues. This plastic response has been shown, however, to be maladaptive in birds and mammals by decreasing cardiac output, decreasing the volume of red blood cells that flows through tissue, and decreasing venous return (Guyton and Richardson, 1961; McGrath and Weil, 1978; Black and Tenney, 1980). An experimental increase of Hct from 42.6% to 65.5 % in dogs while inducing hypoxia ( $PO_2 = 40.5$  mm Hg) resulted in pulmonary vascular resistance increasing by more than 300%. Experimentally increasing Hct alone decreased cardiac output by 50% (McGrath and Weil, 1978).

High elevation hemoglobin specialization provides a precise example of local adaptation as defined by Williams (1966). High elevation specialist genotypes have a competitive advantage under low ambient  $PO_2$  over non-specialist genotypes that have originated from lower elevations (Kawecki and Elbert, 2004). High elevation hemoglobin specialization becomes maladaptive at low elevations. As oxygen binding affinity of the hemoglobin molecule increases, the oxygen-hemoglobin dissociation curve shifts and oxygen becomes more difficult to unload from the active binding site. At high elevations the benefits of increased oxygen affinity outweigh the increased difficulty in releasing oxygen, but this is not the case at lower elevations. Organisms with high elevation hemoglobin are at a disadvantage under normoxic conditions and are thus unable to easily invade lower elevations.

Previous research suggests that the ancestor of the genus *Anairetes* was restricted to high elevations and may have been adapted to hypoxic environments of the high Andes (DuBay and Witt, 2012). If this is the case, *A. reguloides* has most likely secondarily lost its high elevation specialization upon invading lower elevations on the west slope of the Andes. *A. reguloides* occurs above 4000m in the absence of *A. nigrocristatus* but experiences hypoxic stress at these elevations. Because high elevation specialization is maladaptive at low elevations contiguous gene flow with lowland conspecifics limits the ability of *A. reguloides* to redevelop hypoxia resistance. *A. reguloides* is at a physiological disadvantage at high elevations in the presence of *A. nigrocristatus*. Conversely, high elevation specialization limits *A. nigrocristatus* from invading low elevations.

It is important to note that high elevation pressures such as low temperatures and hypoxia may also be driving divergence in the mitochondrial genome. The mitochondrion of eukaryotes is responsible for the production of ATP through oxidative phosphorylation and has been shown to be a target of natural selection in vertebrates living in cold and hypoxic environments (Ehinger et al., 2002; Mishmar et al., 2003; Ruiz-Pesini et al. 2004; Cheviron and Brumfield, 2009). Cheviron and Brumfield (2009) showed mitochondrial haplotype divergence between high and low populations of *Zonotrichia capensis* along a contiguous elevational transect, suggesting functional divergence between elevationally displaced haplogroups. Given the phylogenetic history between *A. reguloides* and *A. nigrocristatus*, the observed 3.5% mitochondrial divergence has most plausibly arose through allopatry. However, upon secondary contact the two mitochondrial haplogroups do not introgress across elevations despite hybridization. It is

possible that the mitochondrial genomes of *A. reguloides* and *A. nigrocristatus* are under selection and are differentially adapted to abiotic pressures of high elevation.

#### 4.2 Restricted gene flow

Divergent evolution of locally adapted taxa has resulted in severely restricted gene flow between *A. reguloides* and *A. nigrocristatus* upon secondary contact. Gene flow between *A. reguloides* and *A. nigrocristatus* is limited to a narrow contact zone at middle elevations with effectively zero introgression out of the Rio Santa Valley and into parental populations. Nuclear haplotype networks suggest disparate retaining of ancestral alleles with limited haplotype sharing between the two species. I would expect through stochastic processes in neutral loci that a low frequency of similar haplotypes might be retained between recently diverged taxa (i.e. incomplete lineage sorting). Coalescent analyses could not, however, differentiate between introgression and incomplete lineage sorting for shared haplotypes. All gene flow estimates statistically overlapped with zero, and I am thus unable to attribute haplotype sharing solely to gene introgression. It is worth noting that I treated all but one shared haplotype as introgressed for cline analyses. This protocol provides a conservative overestimate of cline width. Even so, I observed extremely narrow and dramatic clines. Given the estimated gene flow rates and haplotype networks, at least some of the shared haplotypes are likely the result of incomplete lineage sorting, but with the presence of hybrid individuals I cannot rule out introgression.

Differential adaptation to abiotic pressures at high elevation has resulted in the maintenance of species limits despite reproductive compatibility. Selection for hypoxia resistance, whether in hemoglobin, mtDNA, unidentified markers, or a combination of

loci, favors pure *A. nigrocristatus* genotypes at high elevations and restricts *A. nigrocristatus* alleles from dispersing out of the Rio Santa Valley. Locality three spans high elevations (2900-3800m) on the west side of the Cordillera Negra (Fig. 1) but consists of purely *A. reguloides* phenotypes and genotypes. Based on the observed physiological differences between the two species I would expect to find *A. nigrocristatus* at locality three rather than *A. reguloides*. This exception further supports the physiological disparity between the two species when I take into account geographic barriers. The ridge of the Negra presents a serious east-west barrier to dispersal. The lowest pass in the Negra is ~4200m with barren rock and puna grassland above 4000m. There is also little to no vegetation above 1500m along the northern tip of the range. *Anairetes* are unlikely dispersing over the ridge in either direction or across vast barren stretches of rock or puna grassland. Small, vegetation-restricted passerines such as *Anairetes* would have to disperse from one side to the other along the river valley bottom which is sparsely lined with shrubs. This geographic limitation allows for *A. reguloides* to invade the east side of the Negra from the coast, but restricts *A. nigrocristatus* alleles from dispersing in the opposite direction. The Rio Santa reaches elevations below 1500m before hooking around the northern tip of the Negra and heading towards the coast. Genes adapted to high elevations, whether in pure *A. nigrocristatus* or hybrid individuals, become maladaptive at low elevations and are unlikely to introgress through the low elevation bottleneck that is north of the Negra.

#### 4.3 Mechanism for elevational replacement

Elevational replacement by closely related taxa is well documented in vertebrates throughout montane regions (Terborgh, 1971; Graham, 1983; Sasaki et al., 2005;

Poynton et al., 2007). Previous studies suggest that range expansion of closely related taxa that have become reproductively isolated in allopatry can lead to secondary contact and result in parapatric elevational segregation (Krabbe and Schulenberg, 1997; Cadena, 2007). The mechanisms, however, that promote and maintain stable, vertically displaced distributions are poorly known (Terborgh and Weske, 1975). Our study suggests that differential physiological adaptation to altitude is a mechanism that can promote and maintain elevational segregation upon secondary contact, even in the absence of reproductive isolation. In the valley between the Cordillera Negra and Cordillera Blanca *A. reguloides* and *A. nigrocristatus* segregate elevationally. The valley floor is at an elevation of ~1000m as it hooks around the northern tip of the Negra and increases south to 4000m within a distance of 150km. *A. reguloides* is capable of living at high elevations, as evident from the west side of the Negra, but inhabits only low-mid elevations upon contact with *A. nigrocristatus*. Local alleles provide a competitive advantage at their respective elevations. Individuals with *A. nigrocristatus* alleles are favored at high elevation while individuals with *A. reguloides* alleles are favored lower in the valley. It is worth noting that *A. nigrocristatus* can occur down to ~2200m in the absence of *A. reguloides* (e.g. Field Museum of Natural History specimen 278340 from Acomayo, Huánuco, Peru). The elevational distributions of the two species in allopatry further support the presence of a competitive boundary in the Rio Santa Valley where locally adapted alleles segregate elevationally, restricting gene flow.

## **Acknowledgements**

This study would not have been possible without funding in part by NSF DEB-1146491, NCRN-NIH P20RR018754, the UNM Center for Evolutionary and Theoretical Immunology (CETI), The American Ornithologists' Union, and the Frank M. Chapman Memorial Fund of the American Museum of Natural History. I thank Phred M. Benham, Natalie A. Wright, Matthew R. Jones, Emil Bautista Obispo, Homan Castillo, Andrew B. Johnson, C. Jonathan Schmitt, Donna C. Schmitt, Glenn F. Seeholzer, and Centro de Ornitología y Biodiversidad (CORBIDI; Thomas Valqui, Emil Bautista Obispo, Willy Nanez) for fieldwork and logistical support to obtain key samples for this study. I thank Adam Porter for advice in implementing ClineFit.

## **Literature cited**

- Axelsson, E., N. Smith, H. Sundstrom, S. Berlin, and H. Ellegren. 2004. Male-biased mutation rate and divergence in autosomal, Z-linked and W-linked introns of chicken and turkey. *Molecular Biology and Evolution* 21:1538-1547.
- Bandelt, H., P. Forster, and A. Rohl. 1999. Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* 16:37-48.
- Beall, C. 2007. Two routes to functional adaptation: Tibetan and Andean high-altitude natives. *Proceedings of the National Academy of Sciences of the United States of America* 104:8655-8660.
- Black, C., and S. Tenney. 1980. Oxygen transport during progressive hypoxia in high-altitude and sea-level waterfowl. *Respiration Physiology* 39:217-239.

- Cadena, C. 2007. Testing the role of interspecific competition in the evolutionary origin of elevational zonation: An example with Buarremon brush-finches (Aves, Emberizidae) in the neotropical mountains. *Evolution* 61:1120-1136.
- Chappell, M., J. Hayes, and L. Snyder. 1988. Hemoglobin polymorphisms in deer mice (*Peromyscus maniculatus*), physiology of beta-globin variants and alpha-globin recombinants. *Evolution* 42:681-688.
- Cheviron, Z., and R. Brumfield. 2009. Migration-selection balance and local adaptation of mitochondrial haplotypes in Rufous-collared Sparrow (*Zonotrichia capensis*) along an elevational gradient. *Evolution* 63:1593-1605.
- DuBay, S. G., and C. C. Witt. 2012. An improved phylogeny of Andean tit-tyrants: More characters trump sophisticated analyses. *Molecular Phylogenetics and Evolution*. <http://dx.doi.org/10.1016/j.ympev.2012.04.002>
- Edgar, R. 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 5:1-19.
- Edwards, A. W. F. 1972. Likelihood: an account of the statistical concept of likelihood and its application to scientific inference. Cambridge University Press, Cambridge, UK.
- Ehinger, M., P. Fontanillas, E. Petit, and N. Perrin. 2002. Mitochondrial DNA variation along an altitudinal gradient in the greater white-toothed shrew, *Crocidura russula*. *Molecular Ecology* 11:939-945.
- Felsenstein, J. 1976. Theoretical population genetics of variable selection and migration. *Annual Review of Genetics* 10:253-280.

- Graham, G. 1983. Changes in bar species-diversity along an elevational gradient up the Peruvian Andes. *Journal of Mammalogy* 64:559-571.
- Guyton, A., and T. Richardson. 1961. Effect of hematocrit on venous return. *Circulation Research* 9:157-164.
- Hedrick, P. 1986. Genetic polymorphism in heterogeneous environments- a decade later. *Annual Review of Ecology and Systematics* 17:535-566.
- Hedrick, P., M. Ginevan, and E. Ewing. 1976. Genetic polymorphism in heterogeneous environments. *Annual Review of Ecology and Systematics* 7:1-32.
- Hey, J. 2005. On the number of New World founders: a population genetic portrait of the peopling of the Americas. *Plos Biol.* 3:965–975.
- Hey, J. 2007. Instruction to IM and IMA computer programs.  
<http://genfaculty.rutgers.edu/hey/software>
- Hey, J. 2010. Isolation with migration models for more than two populations. *Molecular Biology and Evolution* 27:905-920.
- Hey, J. 2011. Documentation for IMA2. <http://genfaculty.rutgers.edu/hey/software>
- Hudson, R., and N. Kaplan. 1985. Statistical properties of the number of recombination events in the history of a sample of DNA sequences. *Genetics* 111:147-164.
- Johnson, A. B., McNew, S. M., Graus, M. S., Witt, C. C., 2011. Mitochondrial DNA and meteorological data suggest a Caribbean origin for New Mexico's first Sooty Tern (*Onychoprion fuscatus*). *Western Birds* 42, 233-242.
- Kawecki, T., and D. Ebert. 2004. Conceptual issues in local adaptation. *Ecology Letters* 7:1225-1241.

- Kisdi, E. 2002. Dispersal: Risk spreading versus local adaptation. *American Naturalist* 159:579-596.
- Krabbe, N., & T. S. Schulenberg. 1997. Species limits and natural history of *Scytalopus tapaculos* (Rhinocryptidae), with descriptions of the Ecuadorian taxa, including three new species. *Ornithological Monographs* 48: 47–88.
- Librado, P., and J. Rozas. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25:1451-1452.
- Maddison, D. R., Maddison, W. P., 2005. *MacClade*. Sinauer Associates, Sunderland, MA.
- Maynard Smith, J., and J. Haigh. 1974. The hitch-hiking affect of a favorable gene. *Genetical Research* 23:23-35.
- McCracken, K., C. Barger, and M. Sorenson. 2010. Phylogenetic and structural analysis of the HbA (alpha(A)/beta(A)) and HbD (alpha(D)/beta(A)) hemoglobin genes in two high-altitude waterfowl from the Himalayas and the Andes: Bar-headed goose (*Anser indicus*) and Andean goose (*Chloephaga melanoptera*). *Molecular Phylogenetics and Evolution* 56:649-658.
- McCracken, K., M. Bulgarella, K. Johnson, M. Kuhner, J. Trucco, T. Valqui, R. Wilson, and J. Peters. 2009. Gene flow in the face of countervailing selection: Adaptation to high-altitude hypoxia in the beta A hemoglobin Subunit of Yellow-billed Pintails in the Andes. *Molecular Biology and Evolution* 26:815-827.
- McCracken, K., and R. Wilson. 2011. Gene flow and hybridization between numerically imbalanced populations of two duck species in the Falkland Islands. *Plos One* 6: e23173. doi:10.1371/journal.pone.0023173

- McGrath, R., and J. Weil. 1978. Adverse effects of normovolemic polycythemia and hypoxia on hemodynamics in dogs. *Circulation Research* 43:793-798.
- Mishmar, D., E. Ruiz-Pesini, P. Golik, V. Macaulay, A. Clark, S. Hosseini, M. Brandon, K. Easley, E. Chen, M. Brown, R. Sukernik, A. Olckers, and D. Wallace. 2003. Natural selection shaped regional mtDNA variation in humans. *Proceedings of the National Academy of Sciences of the United States of America* 100:171-176.
- Monge, C., and F. Leonvelarde. 1991. Physiological adaptation to high-altitude oxygen transport in mammals and birds. *Physiological Reviews* 71:1135-1172.
- Porter, A., and E. Levin. 2007. Parallel evolution in sympatric, hybridizing species: performance of *Colias* butterflies on their introduced host plants. *Entomologia Experimentalis Et Applicata* 124:77-99.
- Porter, A., R. Wenger, H. Geiger, A. Scholl, and A. Shapiro. 1997. The *Pontia daplidice-edusa* hybrid zone in northwestern Italy. *Evolution* 51:1561-1573.
- Poynton, J., S. Loader, E. Sherratt, and B. Clarke. 2007. Amphibian diversity in east African biodiversity hotspots: Altitudinal and latitudinal patterns. *Biodiversity and Conservation* 16:1103-1118.
- Pritchard, J., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155:945-959.
- Rezende, E., F. Gomes, C. Ghalambor, G. Russell, and M. Chappell. 2005. An evolutionary frame of work to study physiological adaptation to high altitudes. *Revista Chilena De Historia Natural* 78:323-336.

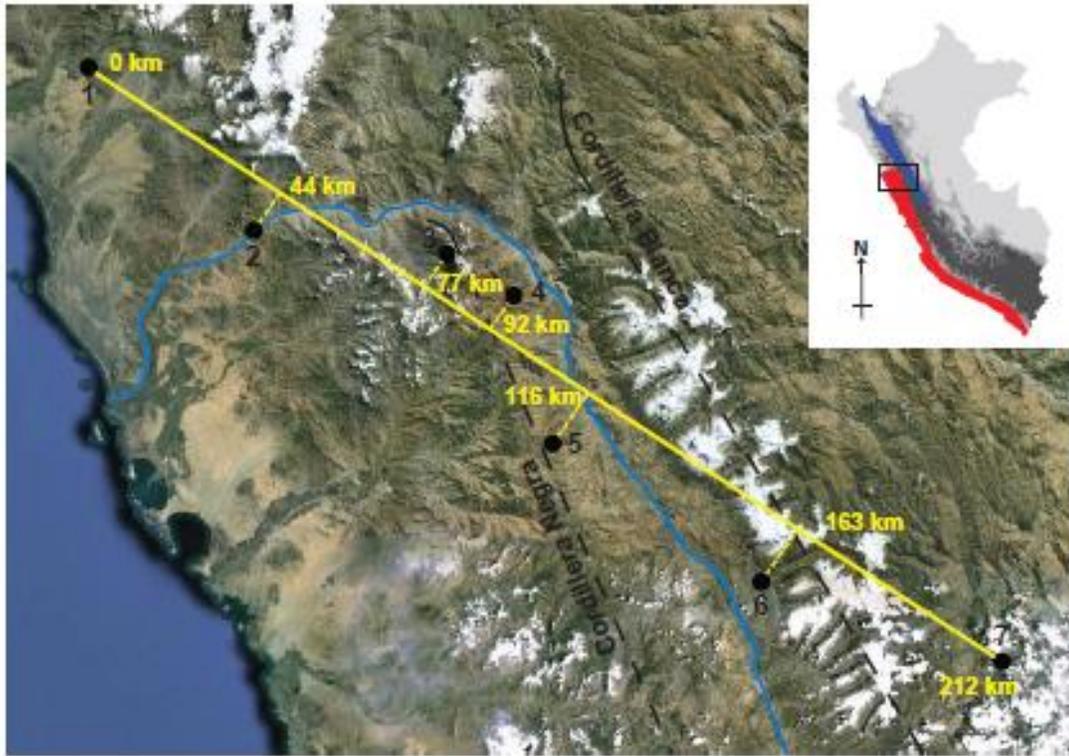
- Ruiz-Pesini, E., D. Mishmar, M. Brandon, V. Procaccio, and D. Wallace. 2004. Effects of purifying and adaptive selection on regional variation in human mtDNA. *Science* 303:223-226.
- Sasaki, K., M.W. Palmer, T. Hayashi. 2005. Influence of climate, elevation, and land use in regional herpetofaunal distribution in Tochigi Prefecture, Japan. *Community Ecology* 6:219-227.
- Schluter, D. 2001. Ecology and the origin of species. *Trends in Ecology & Evolution* 16:372-380.
- Schneider, C., T. Smith, B. Larison, and C. Moritz. 1999. A test of alternative models of diversification in tropical rainforests: Ecological gradients vs. rainforest refugia. *Proceedings of the National Academy of Sciences of the United States of America* 96:13869-13873.
- Schulenberg, T. S., Stotz, D. F., Lane, D. F., O'Neill, J. P., Parker, T. A., III. 2007. *Birds of Peru*. Princeton University Press, New Jersey.
- Sequeira, F., D. Sodre, N. Ferrand, J. Bernardi, I. Sampaio, H. Schneider, and M. Vallinoto. 2011. Hybridization and massive mtDNA unidirectional introgression between the closely related Neotropical toads *Rhinella marina* and *R. schneideri* inferred from mtDNA and nuclear markers. *BMC Evolutionary Biology* 11:264.
- Simmons, P., and A. Lill. 2006. Development of parameters influencing blood oxygen carrying capacity in the welcome swallow and fairy martin. *Comparative Biochemistry and Physiology a-Molecular & Integrative Physiology* 143:459-468.
- Snyder, L. 1981. Deer mouse hemoglobins: Is there genetic adaptation to high-altitude? *Bioscience* 31:299-304.

- Stephens, M., N. Smith, and P. Donnelly. 2001. A new statistical method for haplotype reconstruction from population data. *American Journal of Human Genetics* 68:978-989.
- Storz, J. 2007. Molecular adaptation of hemoglobin function in high-altitude mice. *Comparative Biochemistry and Physiology a-Molecular & Integrative Physiology* 148:S118-S118.
- Storz, J. 2010. Genes for High Altitudes. *Science* 329:40-41.
- Storz, J., and J. Kelly. 2008. Effects of spatially varying selection on nucleotide diversity and linkage disequilibrium: Insights from deer mouse globin genes. *Genetics* 180:367-379.
- Storz, J., A. Runck, H. Moriyama, R. Weber, and A. Fago. 2010a. Genetic differences in hemoglobin function between highland and lowland deer mice. *Journal of Experimental Biology* 213:2565-2574.
- Storz, J., S. Sabatino, F. Hoffmann, E. Gering, H. Moriyama, N. Ferrand, B. Monteiro, and M. Nachman. 2007. The molecular basis of high-altitude adaptation in deer mice. *Plos Genetics* 3:448-459.
- Storz, J., G. Scott, and Z. Cheviron. 2010b. Phenotypic plasticity and genetic adaptation to high-altitude hypoxia in vertebrates. *Journal of Experimental Biology* 213:4125-4136.
- Stotz, D. F., Fitzpatrick, J. W., Parker, T. A., III, Moskovits, D. K., 1996. *Neotropical Birds: Ecology and Conservation*. University of Chicago Press, Chicago, IL.

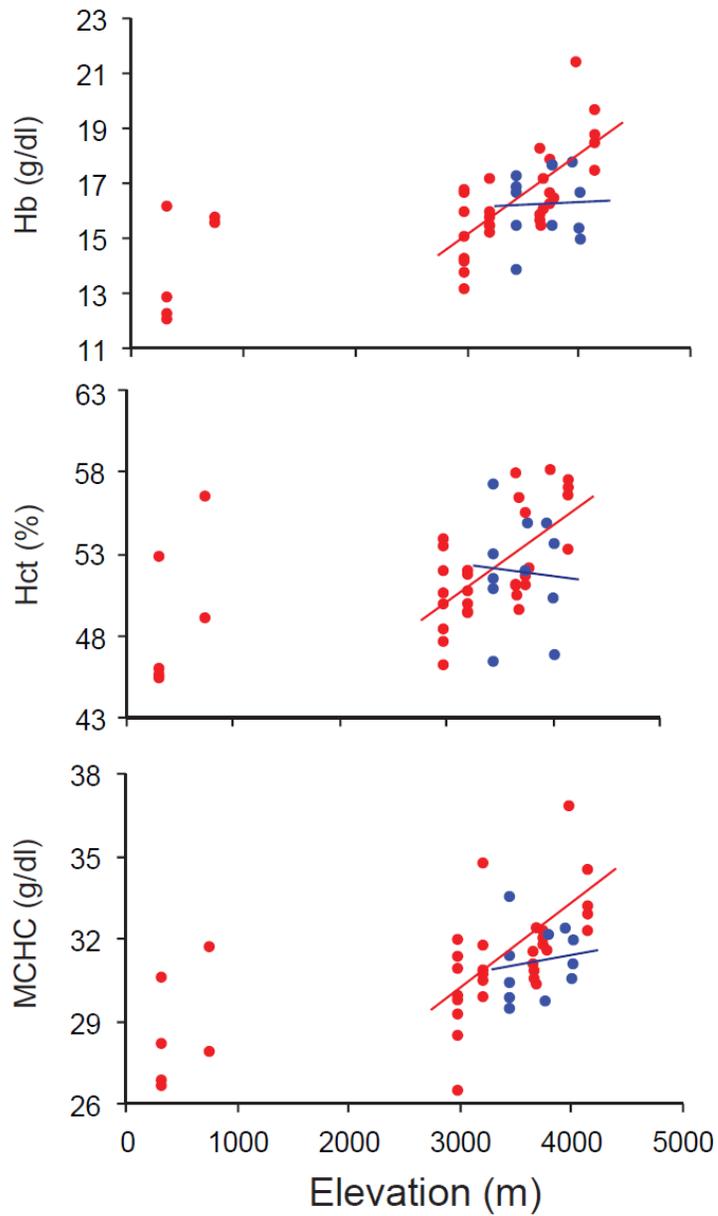
- Szymura, J., and N. Barton. 1986. Genetic analysis of a hybrid zone between the Fire-bellied toads, *Bombina bombina* and *Bombina variegata*, near Cracow in southern. *Evolution* 40:1141-1159.
- Terborgh, J., and J. Weske. 1975. Role of competition in distribution of Andean birds. *Ecology* 56:562-576.
- Turelli, M., N. Barton, and J. Coyne. 2001. Theory and speciation. *Trends in Ecology & Evolution* 16:330-343.
- Terborgh, J. 1971. Distribution on environmental gradients: Theory and a preliminary interpretation of distributional patterns in the avifauna of the Cordillera Vilcabamba, Peru. *Ecology* 52:23-40.
- Via, S. 2001. Sympatric speciation in animals: the ugly duckling grows up. *Trends in Ecology & Evolution* 16:381-390.
- Williams, G. C. 1966. *Adaptation and Natural Selection*. Princeton University Press, Princeton, New Jersey.

**Table 4.** Estimated parameters in the 2 population IMA2 analysis.

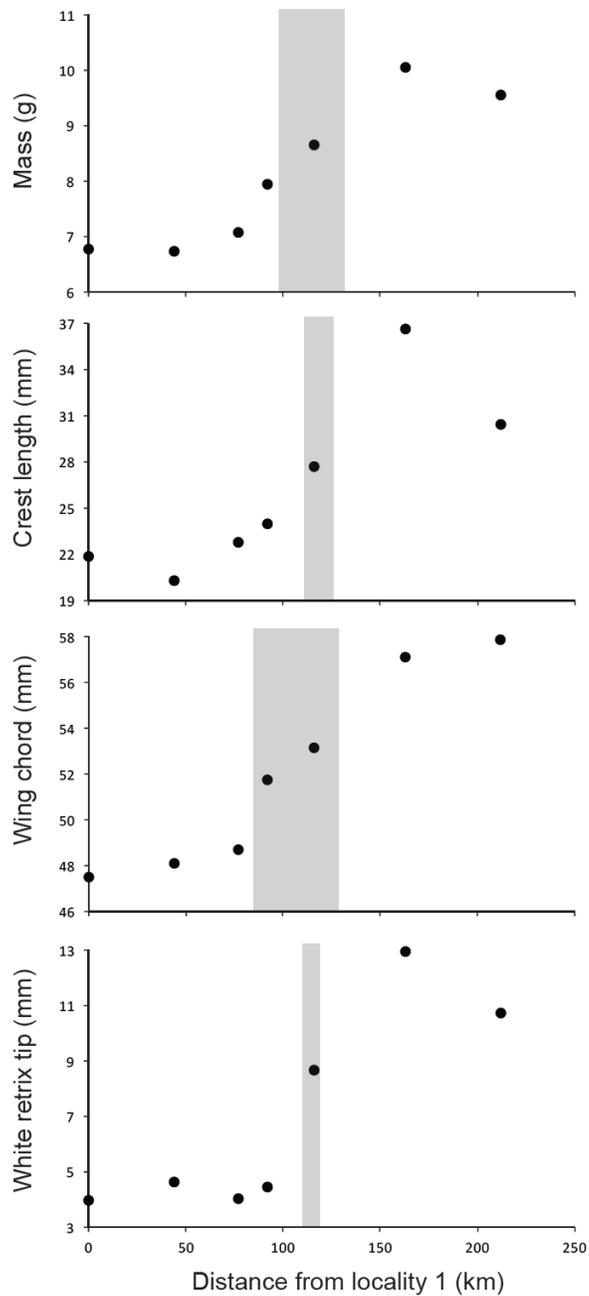
Parameter	Symbol	Definition
Population size parameter ( $4N_e\mu$ )	$\Theta_r$	<i>A. reguloides</i> population
	$\Theta_n$	<i>A. nigrocristatus</i> population
	$\Theta_A$	ancestral population
Time since divergence	$t$	between <i>A. reguloides</i> and <i>A. nigrocristatus</i>
Gene flow ( $M/\mu$ )	$m_{n>r}$	from <i>A. reguloides</i> into <i>A. nigrocristatus</i>
	$m_{r>n}$	from <i>A. nigrocristatus</i> into <i>A. reguloides</i>
Population migration rate ( $2N_eM$ )	$2N_eM_{n>r}$	from <i>A. reguloides</i> into <i>A. nigrocristatus</i>
	$2N_eM_{r>n}$	from <i>A. nigrocristatus</i> into <i>A. reguloides</i>



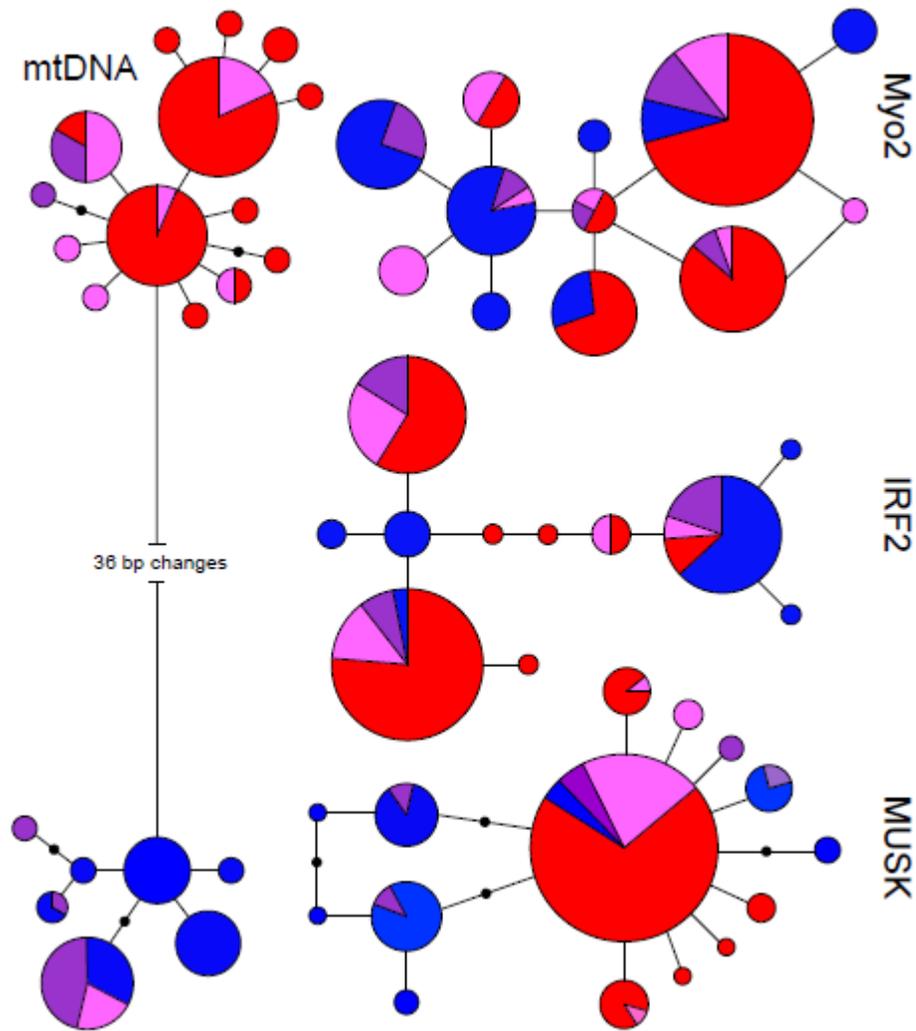
**Figure 6.** Transect across contact zone between *A. reguloides* and *A. nigrocristatus*. Map of Peru in upper right corner. Red shading = range of *A. reguloides*. Blue shading = range of *A. nigrocristatus*. Block dots = sampling localities 1-7. Yellow numbers = distance in km from locality one for each locality. Black dashed lines = ridge line of respective mountain range. The blue line tracks the Santa River.



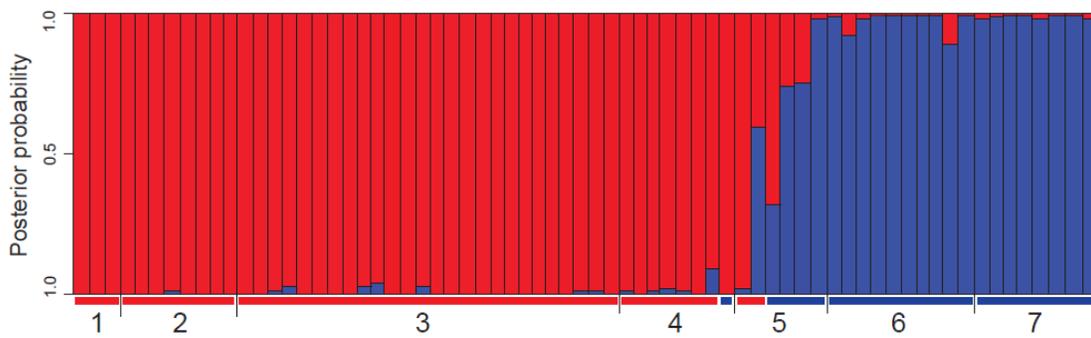
**Figure 7.** Blood-oxygen carrying capacity parameters across elevation. Each point denotes a sampled individual. Red = *A. reguloides*. Blue = *A. nigrocristatus*. Regression lines are plotted for each species above 2900m.



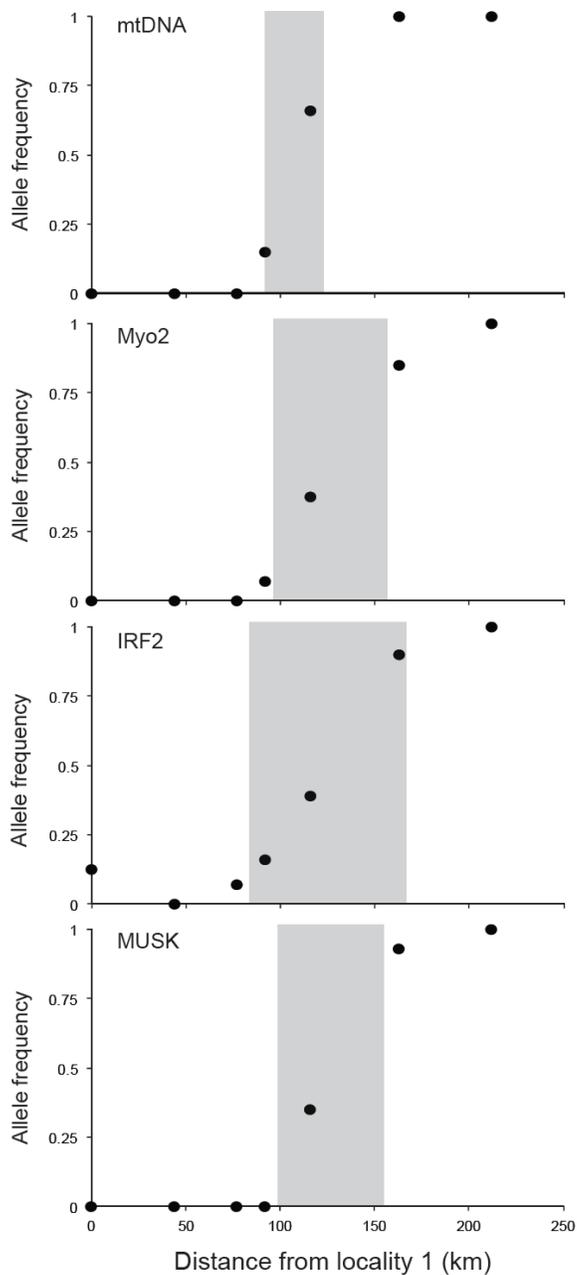
**Figure 8.** Morphometric clines across the contact zone between *A. reguloides* and *A. nigrocristatus*. Each point denotes the average of the trait from each locality (1-7). Shading = cline width positioned around cline center.



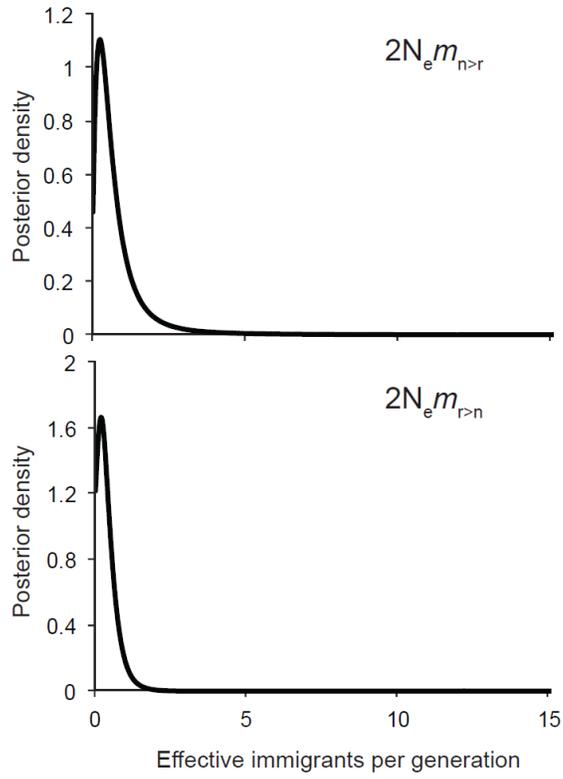
**Figure 9.** Haplotype networks for the four loci. Red = *A. reguloides* phenotype localities (1-3), Blue = *A. nigrocristatus* phenotype localities. Localities 4 and 5 had individuals with intermediate haplotypes and could not be definitively classified as a phenotypic locality for either species. Pink = haplotypes at locality 4. Purple = haplotypes at locality 5.



**Figure 10.** Bayesian probability of species assignment in Structure across transect for the three nuclear loci. Each vertical bar corresponds to a sampled individual. Red = *A. reguloides*. Blue = *A. nigrocristatus*. Numbers correspond with locality along the transect. Horizontal color bars at bottom denote mitochondrial haplotype identity for each individual.



**Figure 11.** Genotype clines across the contact zone between *A. reguloides* and *A. nigrocristatus*. 0% denotes a pure *A. reguloides* locality. 100% denotes a pure *A. nigrocristatus* locality. Shading = cline width positioned around cline center.



**Figure 12.** Posterior densities of population migration rates.

## Conclusions

In this thesis, I have presented two complementary studies that build on the current understanding of speciation and lineage diversification in montane vertebrates. *Anairetes* tit-tyrants have proven to be an informative group for studies of phylogenetics, molecular divergence, ecological divergence, and adaptive evolution. *Anairetes* provides an ideal system for the investigation of diversification and adaptive evolution to abiotic pressures for a number of reasons, including: 1) members show varying degrees of sympatry, allopatry, and parapatry within the genus, 2) some members of the genus are restricted to upper elevations and others span elevational gradients of >4000m, 3) members show varying degrees of habitat specialization and generalization. My first chapter provides valuable insight into phylogenetic methodology by assessing changes over the past decade in both technology and theory. This chapter provides an essential phylogenetic framework to address subsequent questions of adaptive evolution within *Anairetes*. For example, the phylogeny strongly suggests that the ancestor of *Anairetes sensu stricto* inhabited high elevations of the Andes and may have specialized on hypoxic environments. *A. reguloides* has thus secondarily lost high-elevation specialization rather than *A. nigrocristatus* evolving hypoxia resistance.

The second chapter of my thesis highlights the importance of natural selection in driving diversification and promoting biodiversity in montane ecosystems. The evolutionary histories of *A. reguloides* and *A. nigrocristatus* under disparate selection regimes have resulted in differential physiological response to high elevation environments. The differential adaptation between the two species has further resulted in a stable hybrid zone with effectively zero introgression between parental populations.

Despite reproductive compatibility, gene flow is restricted and species limits are maintained. This thesis further supports current adaptive evolutionary theory and is an example of local adaptation maintaining variation and promoting divergence between sister-taxa. Studies of adaptive evolution provide valuable insight into how lineages diverge genetically and ecologically and continue to emphasize the strength of natural selection as a mechanism that drives speciation.