UNIVERSITY OF CENTRAL OKLAHOMA

Edmond, Oklahoma

Jackson College of Graduate Studies

GENETIC ANALYSIS AND ECOLOGICAL NICHE MODELLING OF AN *ARCHILOCHUS* HYBRID ZONE REVEALS DUALITY IN HYBRID ZONE CHARACTERISTICS

A THESIS

SUBMITTED TO THE GRADUATE FACULTY

In partial fulfillment of the requirements

for the degree of

MASTER OF SCIENCE IN BIOLOGY

By

Erica R. Judd

Oklahoma City, Oklahoma

May 2016

GENETIC ANALYSIS AND ECOLOGICAL NICHE MODELLING OF AN *ARCHILOCHUS* HYBRID ZONE REVEALS DUALITY IN HYBRID ZONE CHARACTERISTICS

A THESIS

APPROVED FOR THE DEPARTMENT OF BIOLOGY

May 2016

By Christopher Butler

Dr. Christopher Butler Committee Chairperson

Dr. Michelle Haynie Committee Member

Jenna Hellack

Dr. Jenna Hellack Committee Member

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I would like to thank the University of Central Oklahoma Office of Research and Grants for partially funding this project. Thanks also to the United States Fish and Wildlife Services for permission to band on wildlife refuges, to Brent Ortego in the Texas Parks and Wildlife Department for sending feathers for analysis, to Bob and Martha Sargent for hummingbird banding training and for sending feathers from Alabama, to Eric Judd and Emily Stine-York for field assistance, and to the many hummingbird enthusiasts that invited us to their homes to band.

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES – CONTINUED

ABSTRACT

Two species of hummingbirds (*Archilochus colubris* and *A. alexandri*) whose breeding ranges and timing overlap in the great plains from south Texas to northern Oklahoma were hypothesized to produce hybrids in this area. Proposed hybrids had a gorget color that was intermediate between both parental species. In summer 2010 and 2011, during the breeding season, 200 *Archilochus* individuals were captured and feathers were taken for DNA analysis. The mitochondrial cytochrome oxidase 1 (CO1) gene was sequenced and 3 microsatellite markers were genotyped in order to identify hybrid individuals. Seventy-two unique haplotypes were identified in the sequence data, and phylogenetic trees constructed showed very close relationships among the haplotypes. Based on the difference between field identification and population assignment in Structure, 57 hybrids were identified (95% probability interval > 0.75). Their distribution was modeled using GARP, ENFA, and Maxent and an AUC measure of model accuracy indicated that GARP best represented the distribution of *Archilochus* hybrids (AUC = 0.94). Overall, a broad area of introgression was found to exist from west Texas to central Texas and central Oklahoma. This was farther west than expected because apparent hybrid males have only been documented as far west as Chickasha, OK. Recent expansion of Black-chinned Hummingbirds into Oklahoma suggests hybrid zone transience. However, the center of genetic diversity occurred in the older portions of the hybrid zone (Texas), and environmental variables predicted habitat suitability maps that were consistent with genetic data, which suggests bounded hybrid superiority. This duality indicates that hybrid zones fit different theories at different ages and dynamics. Future research should include multiple *Archilochus* specific microsatellites as well as sequencing additional mitochondrial genes to determine if a genetic cline exists across the US.

Chapter I

INTRODUCTION

Biological Diversity

Biodiversity, the number of species in a given area, is the foundation that supports the interdependent web of life. Systems that are high in biodiversity capture a greater proportion of biologically available resources than systems with a single, dominant species (Cardinale 2011). For example, rivers with a greater diversity of algae species store nitrogen more efficiently than rivers with a single species (Cardinale 2011), grasslands with a higher diversity are more productive and utilize resources more completely (Spehn et al. 2005), and predator diversity in aquatic systems not only affects larval insect diversity in the water, but also affects the diversity of predators in the landscape around the pond (Wesner 2012). Biodiversity is dependent on many factors including temperature, precipitation, latitude, and others. For example, tropical ecosystems, such as the rainforests of the western Amazon, support the most biodiversity (Bass et al. 2010), whereas polar regions, such as Antarctic glaciers, support the least biodiversity (Gaston 2000). Ecological succession provides a timeline by which we can estimate the efficiency of an ecosystem by comparing the available energy flow to biomass and symbiotic relationships (energy exchange) between organisms (Odum 1969). The physical environment, including temperature, precipitation, and soil type, helps determine the rate at which succession proceeds, but the later the succession in a given area, the greater the biodiversity (Odum 1969).

The interrelationship of a species with all the biotic and abiotic factors affecting it is known as a fundamental ecological niche (Grinnell 1917). Genetic adaptation to a specific niche can improve the fitness, or ability to survive and reproduce, of an individual (Orr 2005). For

example, *Escherichia coli*, when exposed to the same temperature for thousands of generations, genetically adapts to that temperature, which improves fitness in later generations (Bennett et al. 1992). Mutations provide an opportunity for organisms to take advantage of resources that otherwise would not have been available to them (Rainey and Travisano 1998). This mechanism of niche differentiation can result in closely related species sharing resources within the same area because of behavioral or morphological variation (Pyke 1982, Olson et al. 1988, Lawler and Morin 1993). Niche specificity increases codependence and ultimately coevolution of organisms, which produces an interconnected web of life (Yoder and Nuismer 2010). Biodiversity ultimately increases efficiency of the ecosystem and can help protect it from violent fluctuations of the physical environment (Odum 1969).

Speciation

Species are the fundamental units used to examine biological diversity (Gill 2007); however, what defines a species has been debated since Darwin (Zink 1996, Avise and Wollenberg 1997). A commonly accepted definition of a species is the Biological Species Concept (BSC, Mayr 1970, Coyne and Orr 2004) which states, "Species are groups of freely interbreeding natural populations that are reproductively isolated from other such groups." Because the BSC emphasizes mate selection as the limiting factor in genetic flow between organisms, it fails to account for geographically isolated populations, interspecies breeding, and the recognition of non-monophyletic groups, and has been repeatedly challenged (e.g., Cracraft 1983, Donoghue 1985, McKitrick and Zink 1988). Preference for a species concept that emphasizes evolutionary history over reproductive relationships has given substance to several versions of the Phylogenetic Species Concept (PSC). One of them states: "A species is the

smallest diagnosable cluster of individual organisms within which there is a parental pattern of ancestry and descent" (Cracraft 1983). A second version states: "We define species as the smallest aggregation of populations (sexual) or lineages (asexual) diagnosable by a unique combination of character states in comparable individuals (semaphoronts)" (Nixon and Wheeler 1990). The Genetic Species Concept, which includes genetic and sexual selection components, defines a species as "a group of genetically compatible interbreeding natural populations that is genetically isolated from other such groups" (Baker and Bradley 2006). Regardless of which species concept is utilized, the mechanism that leads to speciation is likely the same: isolation followed by genetic divergence via natural selection, sexual selection, genetic drift, or a combination of these processes (Zink 1996).

Genetic mutations that produce trait variation within a species are an important component of speciation (Endler 1977). Different environments or resources favor different traits, which results in divergence among populations (Parker 1987). Differentiation can also be due to sexual selection in which traits such as plumage, size, or song traits are enhanced through selective pressures associated with reproduction (Price 1998). Hybridization, or cross-breeding, between divergent populations can be a natural process that may play a role in completion of the speciation process (Grant and Grant 1996). Therefore, hybrid zones are of interest because they provide a model for studying divergence and speciation (Hewitt 1988).

Hybrid Zone Theory

"Hybrid zones" are regions where genetically distinct populations come into contact and interbreed; a resulting cline, or gradient, exists in the genetic composition of populations across the landscape (Barton and Hewitt 1985). Dispersal ability and strength of selection for or against

hybrids influences the width of hybrid zones, while the continuity of habitat and resources affects the internal structure of the genetic cline (Harrison 1993). Thus, not all hybrid zones have the same characteristics (Barton 1979). Hybrid zones can be transient and ephemeral (Rohwer and Wood 1998), or stable and persistent for generations (Barton and Hewitt 1985).

Transient/Ephemeral Hybrid Zones

Transient hybrid zones typically occur because of habitat modification, differences in parental dispersal, or fitness asymmetries between parental groups (Rohwer and Wood 1998). Competitive interaction between parental groups can lead to the local extinction or replacement of one species by another, such as Golden-winged Warblers (*Vermivora chrysoptera*) being replaced by Blue-winged Warblers (*V. cyanoptera)* and their hybrids within the hybrid zone in the eastern US and Canada (Vallender et al. 2007). The hybrid zone in this case has been advancing into the Golden-winged Warbler breeding range for many years (Ficken and Ficken 1968). It is expected that transient hybrid zones either narrow over time due to perfection of reproductive isolation (or speciation), or they broaden over time due to introgressive hybridization and result in fusion of the two types (Endler 1977). Each of these outcomes has a different implication for the direction of speciation.

Stable Hybrid Zones

An example of a stable hybrid zone hypothesis, the dynamic-equilibrium/tension zone, proposes that hybrids have depressed fitness in comparison with parental types. Gene flow from the hybrid zone to the parental zones is inhibited by selection, and hybridization is restricted to a narrow zone between parental populations (Moore 1977). Stability is maintained by a steady influx of parental alleles into the hybrid population that consistently favor reconstitution of parental genomes (Van Den Bussche et al. 1993). Populations in the center of this hybrid zone

should have a disproportionately high number of first generation hybrids compared to other hybrid generations (Van Den Bussche et al. 1993).

Another stable hybrid zone hypothesis is the bounded hybrid superiority model. Hybrids out-compete the parental types in the width of the ecotone (transition area between biomes) where they occur, but they are strongly selected against in the parental portions of the environmental gradient (Short 1972, Moore 1977, Moore and Buchanan 1985, Moore and Price 1993, Pierotti and Annett 1993). In this model, the different hybrid generations breed with each other, producing a broader array of genotypes than possible with repeated crossing of hybrid types with parental types (Baker et al. 1989).

The hybrid-equilibrium model suggests that hybrid zones are stable because all individuals are able to interbreed successfully and there is no selection against either hybrids or the parent organisms (Van Den Bussche et al. 1993). Genetic drift causes differences in parental populations, but selective pressures do not act on differences (no advantages or disadvantages), so the array of genotypes in the center of these populations should follow Hardy-Weinberg expectations (Van Den Bussche et al. 1993).

Tracking and mapping hybrid zones is a labor intensive process using traditional methods. Geographic Information Systems (GIS) are useful when examining hybrid zones because they have the ability to store, map, and analyze spatial data (Swenson 2008). For example, Kohlmann et al. (1988) used abiotic factors to examine the geographic distribution of chromosomal races in a grasshopper and found that precipitation and temperature influenced the distribution of the races. Swenson (2006) used GIS to identify abiotic factors that help maintain avian suture zones (where several hybrid zones cluster on a continent) and concluded that temperature in the Great Plains was predominantly responsible for maintaining the suture zone in

North America. In order to better understand the dynamics of hybrid zones, it is first necessary to identify where they occur and to identify ecogeographical variables (environmental characteristics) that sustain the contact zone (Martínez-Freiría et al. 2008). One approach for determining these factors is to use ecological niche modeling.

Ecological Niche Modeling

The study of why individual species are present in some areas but absent in others has long interested ecologists. Merriam (1898) created a map of life zones (areas with similar plant and animal communities) for the American Southwest in order to explain the "zonation" of life using temperature, elevation, and latitude. Holdridge (1945) created a scheme for classifying the biota of landmasses by environmental variables such as humidity, annual precipitation, and potential evapotranspiration ratios. The term "niche" was first used by Grinnell (1917) to describe the relatively narrow range of environmental conditions within the California chaparral in which the California Thrashers (*Toxostoma redivivum*) were typically found. Hutchinson (1957) suggested that a "fundamental niche" was an *n*-dimensional hypervolume, where each dimension is represented by a factor (temperature, precipitation, elevation, etc.) that might limit the distribution of a species. However, when attempting to calculate the fundamental niche, the practical number of environmental variables that could be measured was limited, the amount of correlated variables was high, and the data were difficult to interpret when every possible feature of the environment was considered (Green 1971). Numerous multivariate statistical approaches have succeeded in identifying a smaller number of factors that influence the niche of species (Green 1971, Pianka 1974, Petraitis 1979).

Ecological niche modeling (ENM) mathematically relates ecogeographical variables (temperature, precipitation, land use/land cover, elevation, etc.) to the presence or absence of a species and then finds similar suitable habitat on a landscape of interest (Hirzel et al. 2002a). Presence data are often museum or herbarium specimens that have georeferenced coordinates (Baker et al. 1998, Funk et al. 1999, Soberón 1999, Ponder et al. 2001, Stockwell and Peterson 2002). They can also be collected from field survey data. Accurate absence data are difficult to obtain, especially in remote or relatively inaccessible areas where modeling would be most useful (Stockwell and Peters 1999, Anderson et al. 2002). In addition, absence data may be misleading because a species may not be recorded in a given locality due to true absence or because it was not detected (see Hirzel et al. 2002a). The ecogeographical variables utilized in ENM do not cover the entire range of possible ecological niche dimensions (Hutchinson 1957), but the currently available digital maps provide many of the dimensions that influence species distribution (Grinnell 1917, Root 1988, Brown and Lomolino 1998). A model of the species distribution is calculated and then projected onto a map of the study area (Chen and Peterson 1999, Peterson and Vieglais 2001). Numerous ENM approaches exist including BIOCLIM (Santika and Hutchinson 2009), DOMAIN (Allouche et al. 2008), General Additive Models (Gallego et al. 2004), CART (Tognelli et al. 2009), and General Linear Models (Roura-Pascual et al. 2009). However, the three most commonly used models in the literature are Genetic Algorithm for Rule-set Production, Ecological Niche Factor Analysis, and Maximum Entropy.

GARP

Genetic Algorithm for Rule-set Production (GARP, Stockwell and Peters 1999) uses a heuristic (loosely defined rules) approach to identify factors correlated with species presence/absence (Anderson et al. 2003). Because of the random processes used to generate a

niche-model within GARP, model runs may result in slightly different outputs despite using the same input data (Anderson et al. 2003). A test of model performance was developed to find the best model out of many using a confusion matrix (Anderson et al. 2003). Models with high omission error (false negatives) are eliminated, and remaining models are filtered based on an optimal balance between overfitting and overprediction (Anderson et al. 2003). The model can be projected onto the study area to create a map of the study species distribution with probability ranges between 0 and 1. GARP has been used to model a wide variety of organisms including birds (Bergen et al. 2007), mammals (Rubin et al. 2009), reptiles (García 2006), amphibians (Chen and Bi 2007), plants (Sánchez-Flores 2007), insects (Tognelli et al. 2009), and even diseases (Adhikari et al. 2009). Some authors suggest that absence data can be misleading and that "false absences" potentially could skew the results of the model (Hirzel et al. 2002a). Despite this potential bias, GARP appears to be an effective method for predicting presence/absence of a species on the landscape (Tsoar et al. 2007).

ENFA

An alternative approach is to use presence-only data to predict the species range using an Ecological Niche Factor Analysis (ENFA). This factor analysis compares the ecogeographical variable (EGV) values of the species presence points to the EGV values of the whole area of interest and uses the differences to classify each cell according to habitat suitability for the species (Hirzel et al. 2002a). The EGV values can differ with respect to their mean and variance. ENFA summarizes multiple EGVs into a few independent factors that have specific biological significance (Hirzel et al. 2002a). The first factor is "marginality," which maximizes the difference in environmental conditions between the species niche and the general study area. Subsequent factors are "specialization" factors that are created by computing the direction that

maximizes the remaining variance between the study area and the locations where the species are found. The robustness of the model is assessed using a continuous Boyce index (see Hirzel et al. 2006 for a description of this method), which can range from 0 to 1, with models closer to 1 having the best fit. A habitat suitability map can be computed to display the predicted distribution of the organism with probability ranging from 1 to 100 (Hirzel et al. 2002a). ENFA also has been used to model the distribution of a wide range of species including birds (Rivera et al. 2008), mammals (Sattler et al. 2007), reptiles (Santos et al. 2006), amphibians (Soares and Brito 2007), plants (Zaniewski et al. 2002, Butler et al. 2011), and arthropods (Estrada-Peña and Venzal 2007). ENFA may be particularly useful when examining the distributions of organisms at the edge of their range (Braunisch et al. 2008).

MaxEnt

Jaynes (1957) wrote, "Information theory provides a constructive criterion for setting up probability distributions on the basis of partial knowledge, and leads to a type of statistical inference which is called the maximum-entropy estimate." This maximum-entropy (MaxEnt) estimate can be used to model the distribution of organisms given presence localities and environmental covariates (Phillips et al. 2006, Phillips and Dudík 2008). MaxEnt estimates a species' distribution by finding the distribution with maximum entropy given moment (mean, variance) constraints (Elith et al. 2011). MaxEnt starts by assuming that the probability is perfectly uniform in geographic space and moves away from this distribution only to the extent that it is forced to by the constraints (Phillips et al. 2004). It outputs a probability distribution that sums to 1 and represents the relative probability of observing the species in each cell. MaxEnt has been used to predict the ranges of a large number of species including birds (Phillips and Dudik 2008), plants (Butler et al. 2012), mammals (Papeş and Gaubert 2007), reptiles

(Phillips and Dudik 2008), and insects (Ward 2007). Although originally created to analyze presence/absence data, it is possible to run MaxEnt with presence-only data (Phillips and Dudík 2008). It is particularly useful when trying to create a model with a limited number of known locations for the species of interest (Papeş and Gaubert 2007).

Application to Hybrid Zones

Recent studies have used ENM to relate environmental factors to the presence or absence of hybrids, making it possible to predict the extent of a hybrid zone with less labor (Kohlmann et al. 1988, Cicero 2004). Kohlmann et al. (1988) used BIOCLIM (a program that provides estimates of bioclimatic parameters - see Nix 1986 for more details) to examine the distribution of four chromosomal races in the Australian grasshopper (*Caledia captiva*). Cicero (2004) also used BIOCLIM in conjunction with DIVA-GIS to correlate bioclimatic factors with the amount of gene flow between the Oak Titmouse (*Baeolophus inornatus*) and the Juniper Titmouse (*B. ridgwayi*) in northern California. Swenson (2006) used GARP to examine an avian suture zone in the Great Plains and to identify whether climate or topography influenced the location of the suture zone. Martínez-Freiría et al. (2008) used MaxEnt to identify the ecogeographical variables (not just bioclimatic variables, but elevation, habitat, etc.) associated with contact zones of three viper species (*Vipera aspis*, *V. latastei*, and *V. seoanei*) in northern Spain. Chatfield et al. (2010) used MaxEnt to examine the climate factors that affect the hybrid zone location in the southern Appalachian Mountains of three salamander species (*Plethodon jordani*, *P. metcalfi,* and *P. teyahalee*). MaxEnt also was utilized with two peociliid fishes (*Xiphophorus birchmanni* and *X. malinche*) to determine the environmental factors that affect hybrid zone structure (Culumber et al. 2012).

Great Plains Suture Zones

The Great Plains is a good region to study hybridization because of the ecological transition from the drier Rocky Mountain region to the wetter Gulf Coastal Plain. A number of eastern bird species reach their westernmost range and a number of western bird species reach their easternmost range in this region (Rising 1983a, Swenson and Howard 2005). For example, the eastern Tufted Titmouse (*B. bicolor*) meets its southwestern ecological counterpart, the Black-crested Titmouse (*B. atricristatus*), in the Great Plains (Curry and Patten 2014). The Eastern Meadowlark (*Sturnella magna*) replaces the Western Meadowlark (*S. neglecta)* along a suture zone in the Great Plains (Rising 1983a). Additional examples of ecological counterparts meeting in the Great Plains include: Yellow-shafted and Red-shafted varieties of the Northern Flicker (*Colaptes auratus*, Short 1965), Baltimore and Bullock's Orioles (*Icterus galbula* and *I. bullockii*, Sibley and Short 1964), Indigo and Lazuli Buntings (*Passerina cyanea* and *P. amoena*, Sibley and Short 1959), Rose-breasted and Black-headed Grosbeaks (*Pheucticus ludovicianus* and *P. melanocephalus*, West 1962), and Eastern and Spotted Towhees (*Piplo erythrophthalmus* and *P. maculatus*, Sibley and West 1959). Several studies have been conducted on hybridizing bird species in the Great Plains (Rising 1970, 1983b, 1996, Anderson and Daugherty 1974, Kroodsma 1974a, 1974b, 1975, Emlen et al. 1975, Moore and Buchanan 1985, Moore and Koenig 1986, Moore et al. 1991, Moore and Price 1993), but hummingbird contact zones have largely been neglected, although two species have breeding ranges that overlap in the Great Plains.

Archilochus Hybrid Zone

Ruby-throated Hummingbirds (*Archilochus colubris*) breed in the eastern United States and Canada (Figure 1, Robinson et al. 1996). According to the 1997-2001 Oklahoma Breeding Bird Atlas, Ruby-throated Hummingbirds were confirmed as breeders throughout eastern Oklahoma and in southwestern Oklahoma, but were not found in northwestern Oklahoma and the Panhandle (Reinking 2004). Black-chinned Hummingbirds (*A. alexandri*) breed in the western United States from central Texas and northern Mexico west to northern Baja California and north to interior southern British Columbia (Figure 1, Baltosser and Russell 2000). Although many published range maps do not show Black-chinned Hummingbirds breeding in Oklahoma (Baltosser and Russell 2000, Howell 2002), the 1997-2001 Oklahoma Breeding Bird Atlas confirmed Black-chinned Hummingbirds as breeders in Comanche County and Cimarron County and listed them as probable breeders in Stephens County (Reinking 2004). Additionally, Black-chinned Hummingbirds have recently been documented breeding in Grady County, Oklahoma (Butler et al. 2007) and in southwestern Kansas (Thompson et al. 2011), indicating a possible northern shift in their breeding range.

According to NatureServe's online database (www.natureserve.org), breeding ranges of *A. colubris* and *A. alexandri* overlap in much of central Texas (Figure 1). An *Archilochus* hybrid zone is hypothesized to exist in Oklahoma and Texas because of an accumulation of sightings of males that have a different gorget (throat feathers) color than either of the breeding species in the area (Judd et al. 2011).

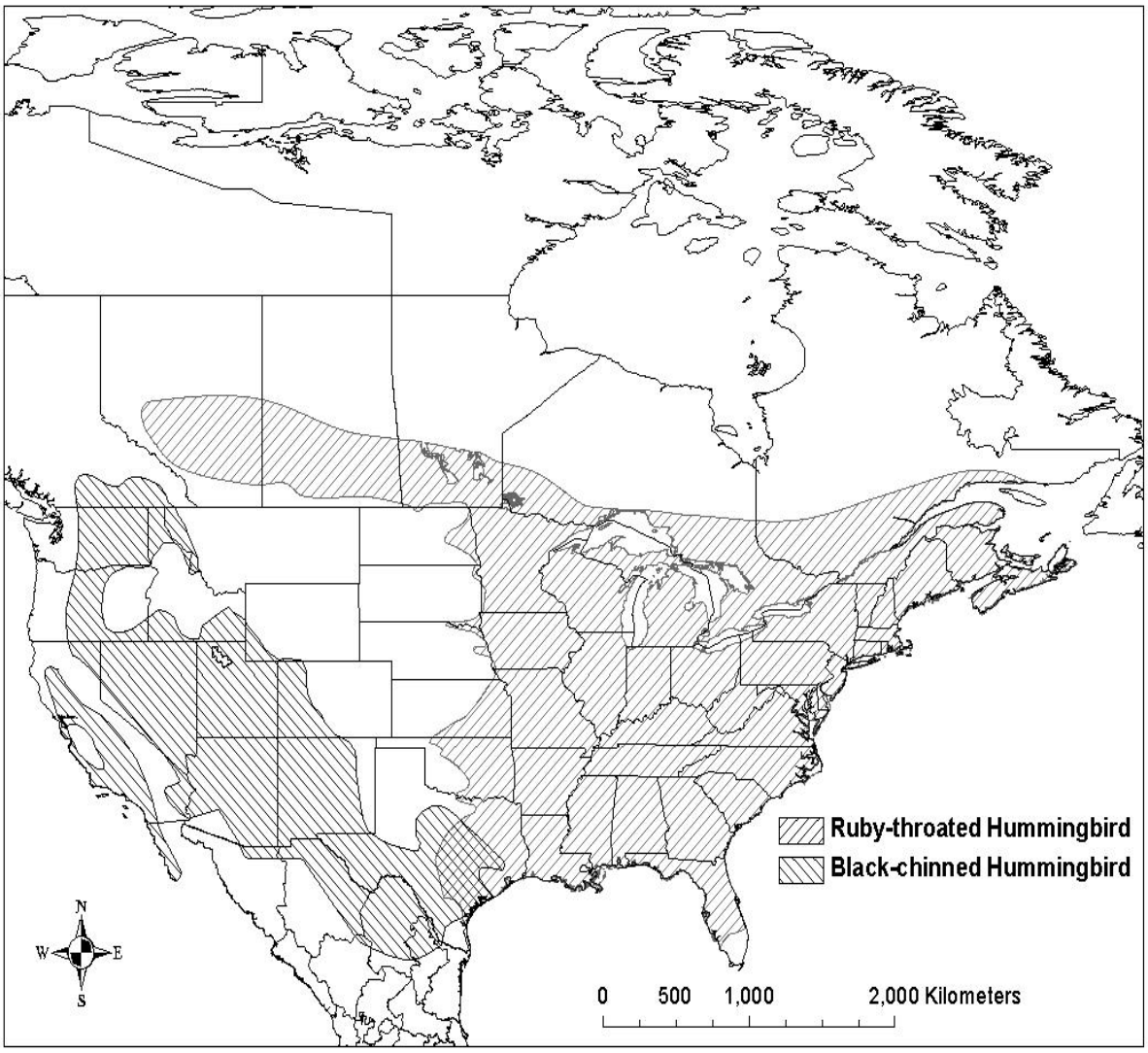

Figure 1: The breeding range of the Black-chinned Hummingbird and Ruby-throated Hummingbird. This map was created using data provided by BirdLife International and NatureServe (2012).

Archilochus Morphology

Ruby-throated Hummingbird males have a red gorget that extends up to their bill (Figure 2). Their backs are emerald green and they have black primaries and rectrices. Their bellies are white with buff-colored flanks. Black-chinned Hummingbird males have a black upper gorget with a band of purple across the lower gorget (Figure 3). Their backs are yellow- green and they have black primaries and rectrices. Their bellies are greyish white with no buff color on their flanks. Hatch year males that are still on their parental breeding grounds may or may not produce a few gorget feathers. The extent of white on the rectrices, as well as primary feather shape, separates them from female *Archilochus* (Pyle 1997).

Female Ruby-throated Hummingbirds (Figure 4) are difficult to distinguish from female Black-chinned Hummingbirds (Figure 5) because of their similar plumages and overlapping measurements. Wing length for Ruby-throated Hummingbird females ranges from 43.4 - 47.3 mm, whereas Black-chinned Hummingbird female wing length ranges from 43.5 - 48.7 mm. Culmen (bill) length of Ruby-throated Hummingbird females ranges from 15.2 - 19.0 mm, whereas bill length of Black-chinned Hummingbird females ranges from 17.9 - 22.9 mm. The tenth (outer) primary feather on adult female Ruby-throated Hummingbirds (Figure 6) is relatively straight and pointed, and the same feather is curved and somewhat spatulate on the adult female Black-chinned Hummingbird (Figure 7). Close examination of the sixth (counting inward) flight feather shape is necessary for identification of females and juveniles. For example, female Ruby-throated Hummingbirds have a pointed sixth primary feather (Figure 6), and female Black-chinned Hummingbirds have a more rounded and wide sixth primary (Figure 7). However, it is not possible to identify female hybrids based on morphological characteristics (Pyle 1997).

Figure 2: Male Ruby-throated Hummingbird

Figure 4: Female Ruby-throated Hummingbird

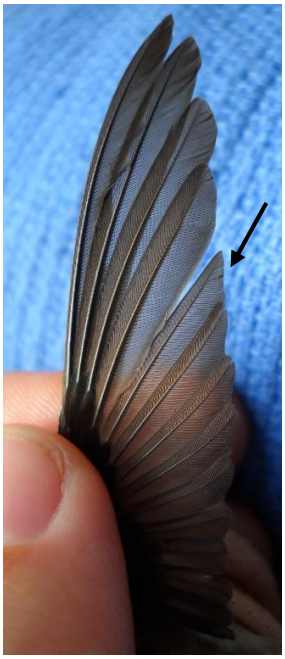

Figure 6: Female Ruby-throated Hummingbird primaries. Arrow indicates 6th primary feather.

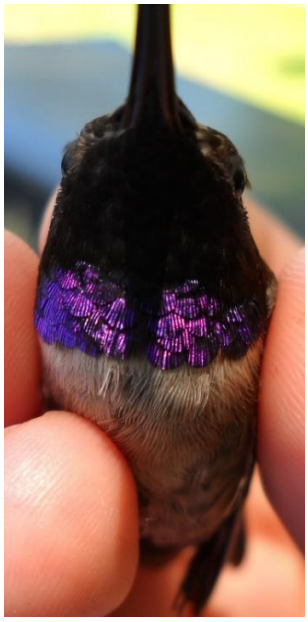

Figure 3: Male Black-chinned Hummingbird

Figure 5: Female Black-chinned Hummingbird

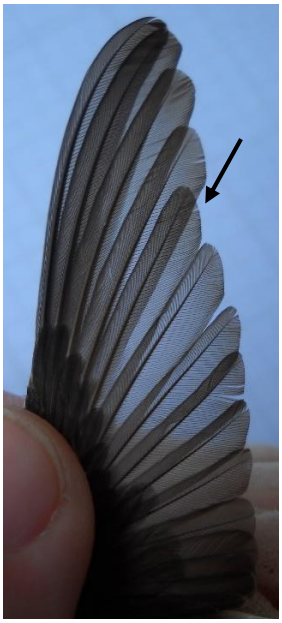

Figure 7: Female Black-chinned Hummingbird primaries. Arrow indicates 6th primary feather.

Possible hybrids of these two species have been observed in central Oklahoma (Oklahoma County, Vacin 1969) and along the Red River in Grayson County, Texas (Pulich 1988). The reported Oklahoma hybrid was male and had purple iridescent feathers with a reddish tint in a band on the lower portion of the gorget (Vacin 1969). The reported Texas hybrid was found dead and was not described (Pulich 1988).

Figure 8: Apparent male hybrid from North Central Texas. Note the pink, iridescent gorget feathers that form a wide band across the front of the neck. Black feathers comprise about 30% of the upper gorget in this case.

Figure 9: Apparent male hybrid from Southwest Texas. Note the pink iridescent band in gorget with black feathers above. Photo courtesy of Brent Ortego.

Twelve *Archilochus* hybrids were banded at four locations in central and southwestern Oklahoma between 2006 and 2009 (Judd et al. 2011). These hybrids were all males and were identified primarily by the pink gorget color (Figures 8 and 9). It is possible that hybrids exist that have genetic introgression without obvious phenotypic variation (e.g., female *Archilochus* hybrids). This complicates identification of hybrids and necessitates the use of molecular techniques to identify individuals with genetic admixture.

Background Work in Hummingbird Phylogeny

The genus *Archilochus* belongs to the Trochilidae family, which includes subfamilies Phaethornithinae (hermits) and Trochilinae (non-hermits). An expanded informal taxonomy as well as a formal phylogenetic taxonomy based on PhyloCode recommended nine clades within Trochilidae, separating the family into topazes, hermits, mangoes, coquettes, brilliants, *Patagona*, mountain gems, bees (including *Archilochus sp*.), and emeralds (McGuire et al. 2009). Using phylogenetic estimates to examine the biogeography of 151 species of hummingbirds, McGuire et al. (2007) determined that all of the principal clades of hummingbirds, except mountain gems and bees, originated in the lowlands of South America. Additionally, most of the hummingbirds present in North America were the result of seasonal invasions from South America (McGuire et al. 2007).

Bleiweiss (1998) used DNA-DNA hybridization to determine large-scale phylogenetic structure in the Trochilidae using 26 species of hummingbirds. Their results indicated that the hermit and non-hermit subfamilies began diverging approximately 17 million years ago (Bleiweiss 1998). Rapid divergence of the hummingbirds into two Andean clades (the coquettes

and brilliants) and one Central and North American clade (emeralds, mountain gems, bees, and *Patagona*) occurred about 12 million years ago (Bleiweiss 1998).

In order to confirm the sequential radiation of Trochilidae, Chubb (2004) used ZENK sequences to reconstruct gene trees containing the major lineages of each group. Chubb (2004) determined that there was evidence that Trochilidae diverged sequentially starting with the hermits, to the mangoes, to the coquettes, to the emeralds, and finally ending in a sister relationship between the mountain-gems and the bees (including *Archilochus*). McGuire et al. (2014) presented an expanded multilocus phylogenetic analysis of 284 hummingbird species, and added that topazes had strong support as a sister taxon of the hermits instead of within a Trochilidae subfamily. This placed the topazes before hermits in the sequence of hummingbird radiation; however, the timing of all other hummingbird radiations was consistent with Bleiweiss (1998) and Chubb (2004).

Genetic Identification of Hybrids

Much of the background work in phylogenetics was made possible by utilizing DNA barcoding primers. Hebert et al. (2004) sequenced the mitochondrial cytochrome oxidase I (COI) gene in order to identify unique barcodes (or single nucleotide polymorphisms) within 260 species of North American birds. Their goal was to establish a basis for species identification of birds in order to clarify avian phylogenetic relationships (Hebert et al. 2004). Although the ability of DNA barcoding to identify sister species correctly has been called into question (Johnson and Cicero 2004), its effectiveness has been demonstrated in multiple studies (see Hebert et al. 2003, 2004, Tavares and Baker 2008). Tavares and Baker (2008) sequenced the COI gene in several sister species pairs, including *A. alexandri* and *A. colubris*, and found five

loci that were fixed for each species. *Archilochus colubris* exhibited a barcode type of CACAA, and *A. alexandri* possessed a barcode type of TCTGG (Tavares and Baker 2008).

Mitochondrial DNA is useful when identifying hybrids because it is inherited only from the maternal lineage (Sato and Sato 2012) making it possible to identify the maternal species. Seneviratne et al. (2012) used morphometric data combined with molecular markers to analyze a sapsucker (*Sphyrapicus*) hybrid zone in North America. Rohwer et al. (2001) also used plumage characteristics and mtDNA to analyze two of the three hybrid zones between Townsend's Warblers (*Setophaga townsendi*) and Hermit Warblers (*S. occidentalis*).

Morphological characters and mtDNA sometimes lack sufficient information to correctly identify hybrids, so inclusion of multiple nuclear microsatellite markers may be beneficial. Smith et al. (2013) used mtDNA and microsatellite markers to characterize a hybrid zone between Australian tree frogs (*Litoria ewingii* and *L. paraewingi*). Miraldo et al. (2013) also used mtDNA and microsatellite markers to measure the diffusion of genes between two lineages of ocellated lizard (*Lacerta lepida*) in their southeastern Iberia hybrid zone. Microsatellites and mtDNA, combined with morphological characteristics, will be used to identify *Archilochus* hybrids.

Study Purpose

There were four goals in this study, the first of which was to identify hybrids using genetics. The second goal was to determine the extent of the *Archilochus* hybrid zone with ecological niche modeling. The third goal was to compare the accuracy of the resulting habitat suitability maps. The final goal was to use the most accurate habitat suitability map to determine which hybrid zone hypothesis best describes the *Archilochus* hybrid zone.

Chapter II

MATERIALS AND METHODS

Study Sites

ORNIS (Constable et al. 2010, Guralnick and Constable 2010) was used to access location data (15 May to 15 July, 1950-2000) for Ruby-throated Hummingbirds and Black-chinned Hummingbirds. Dates were chosen during the breeding season that did not coincide with migration so that genetic samples would represent the area of interest. Quantum GIS Desktop v. 2.6.1 (QGIS Development Team 2014) was used to overlay locations on a map and determine where to concentrate sampling efforts for this study. Bird listservs (OKBirds, TexBirds, and TX-Hummer), U.S. Fish and Wildlife Services, and Texas Parks and Wildlife Department were contacted to find locations to band in Oklahoma and Texas.

From 15 May to 15 July 2010 and 2011, 199 *Archilochu*s DNA samples were collected from within and outside the breeding range overlap in Oklahoma and Texas (Figure 10 and Appendix A). The samples obtained outside the breeding range overlap were used to determine pure parental haplotypes. Hummingbirds were captured with nectar feeders inside wire cage traps or Russell traps (Russell and Russell 2001). They were banded with Bird Banding Lab identification numbers (Gustafson et al. 1997, Federal permit # 23357).

Mass, wing length, bill length, age, and sex were recorded, and pictures were taken of the head, gorget, wing, and rectrices. Species identification of after hatch year (AHY) males was done by gorget color, and species identification of AHY females was done by a combination of bill length, wing length, and shape of the sixth primary feather (Pyle 1997). Hatch year males and females were identified with a combination of bill corrugation, bill length, wing length, outer

primary shape, and the shape of the end of the sixth primary feather (Pyle 1997). Shape of the sixth primary was the determining factor when other observations were inconclusive.

Blood was obtained by pricking the brachial vein with a fine-gauge needle and allowing the blood to pool on the surface of the skin. It was collected with capillary tubes (Roy et al. 1998) and was stored in lysis buffer (White and Densmore 1992). The outer two or inner two rectrices were collected if blood could not be obtained and were stored in paper envelopes to preserve the DNA. In most cases, blood was difficult to obtain and required prolonged handling of the subject, so it was quickly determined that only feathers would be obtained in order to decrease handling time. Bob Sargent provided Ruby-throated Hummingbird parental types from Alabama (samples 1-10 in Appendix A). Black-chinned Hummingbird parental types were collected from Christoval, TX (samples 108-117, Appendix A).

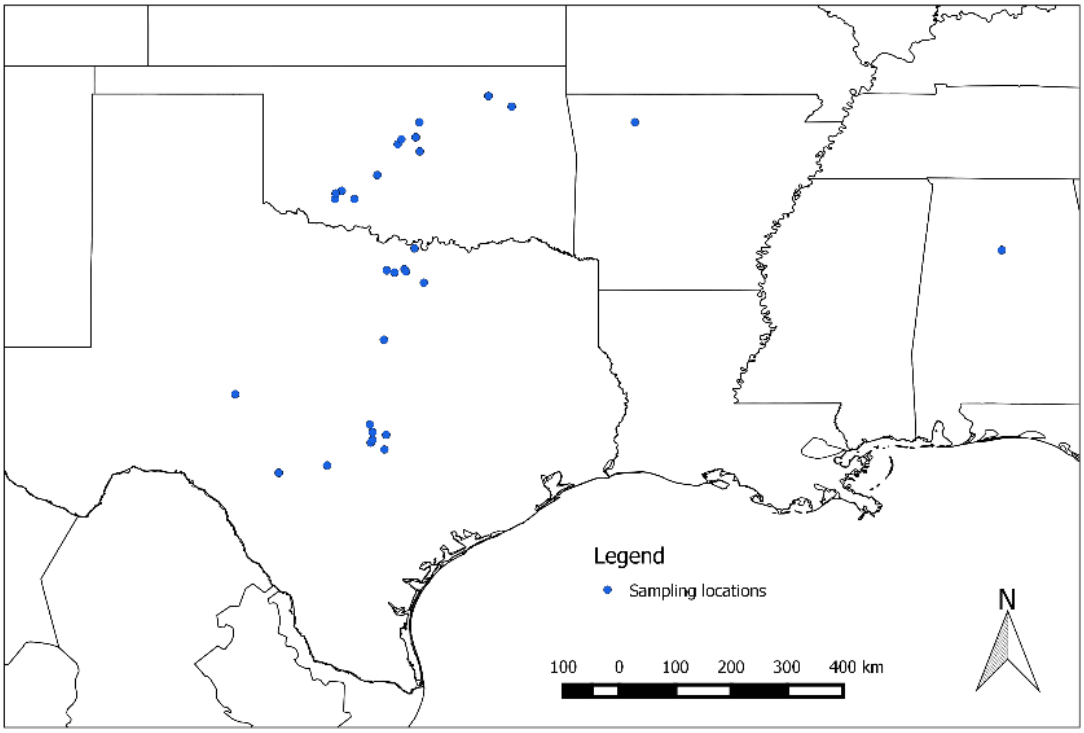

Figure 10: DNA sampling localities. DNA samples were obtained from thirty sites in Oklahoma, Texas, Arkansas, and Alabama from 15 May – 15 July 2010 and 2011.

Lab Work

Extraction

Extractions were performed with a DNeasy blood and tissue kit (Qiagen, Valencia, CA). Blood samples were extracted according to the DNeasy protocol except that 100 µL of dilute blood in lysis buffer (5-30 µL / 1000 µL) was used. The "plucked body contour feathers" DNA extraction protocol (Bush et al. 2005) was used with both rectrices quills in one reaction with the recommended DNeasy blood and tissue kit.

Sequencing

The mitochondrial cytochrome oxidase I (COI) gene was amplified using barcoding primers LTyr and COI907aH2 (Tavares and Baker 2008). PCR amplification was performed in a 25 µL reaction mixture with 1.25 µL LTyr, 1.25 µL COI907aH2, 0.25 µL GoTaq Flexi DNA Polymerase (Promega, Madison, WI), 12.5 µL Premix D (FailSafe PCR system, Epicentre, Madison, WI), 0.5-3.0 µL extracted DNA (5-20 ng), and ddH2O to bring total volume to 25 µL. Thermal cycle conditions were 36 cycles of 94°C for 40 seconds, 53°C for 40 seconds, and 72°C for 1 minute, with an initial denaturation of 94°C for 5 minutes and a final extension of 72°C for 7 minutes.

Amplified fragments were cycle sequenced using primers obtained from Tavares and Baker (2008), and included LTyr (forward) and CO1748Ht (reverse). Post cycle sequencing products were purified using Performa® DTR Gel Filtration Cartridges (Edge BioSystems, Gaithersburg, MD). Purified products were sequenced in both the forward and reverse directions on an ABI3130 (Applied Biosystems, Foster City, CA) using the manufacturers' protocol for cycle sequencing.

Genotyping

Jamaican Streamertail Hummingbird (*Trochilus polytmus*) microsatellite primers (Lance et al. 2008) were utilized because of the scarcity of published *Archilochus* microsatellite primers. Twelve of the primers published by Lance et al. (2008) were optimized and showed variation; however, the small volume of extraction product that remained after sequencing limited the analysis to only three loci. PCR reaction optimization was necessary for each locus because lower quality feather DNA was used instead of the pectoral muscle used in Lance et al. (2008).

PCR amplification of the Tro 2 locus was performed in a 15.0 µL reaction mixture with 7.70 µL ddH_2O, 1.0 µL $MgCl_2$ (Promega), 1.5 µL 10x PCR Buffer (Promega), 1.2 µL dNTP mix (Promega), 0.2 µL each of Tro 2 forward and reverse, 0.2 µL GoTaq® Flexi DNA Polymerase (Promega), and 3 µL extracted DNA.

PCR amplification of the Tro 4 locus was performed in a 15.0 µL reaction mixture with 7.76 µL ddH2O, 1.2 µL $MgCl_2$, 1.5 µL 10x PCR Buffer, 1.2 µL dNTP mix , 0.12 µL each of Tro 4 forward and reverse, 0.1 µL GoTaq® Flexi DNA Polymerase , and 3 µL extracted DNA.

PCR amplification of the Tro 10 locus was performed in a 15.0 µL reaction mixture with 7.96 µL ddH2O, 1.0 µL $MgCl_2$, 1.5 µL 10x PCR Buffer, 1.2 µL dNTP , 0.12 µL each of Tro 10 forward and reverse, 0.1 µL GoTaq® Flexi DNA Polymerase , and 3 µL extracted DNA.

A touchdown thermal profile was used that included a 10°C span in annealing temperature ranging from 55°C - 65°C. Cycling parameters were 20 cycles of 95°C for 30 seconds, 65°C (decreased by 5°C each cycle) for 30 seconds, and 72°C for 30 seconds; and 20 cycles of 95°C for 30 seconds, 55°C for 30 seconds, and 72°C for 30 seconds, with an initial denaturation of 95°C for 2 minutes and a final extension of 72°C for 10 minutes.

Amplification products were run on an ABI3130 Genetic Analyzer with GeneScan 500 ROX size standard (Applied Biosystems) according to manufacturer's instructions. Microsatellite allele calls were determined using GeneMapper Software v. 3.7 (Applied Biosystems).

Phylogenetic Tree Construction

Sequences were checked for ambiguities, edited, and trimmed using Sequencher v. 4.10.1 (GeneCodes Corp., Ann Arbor, MI). Alignments were done in MEGA v. 6 (Tamura et al. 2013). Sequences were aligned using the MUSCLE algorithm (Edgar 2004), which is incorporated into MEGA 6. The CLUSTALW (Thompson et al. 1994) alignment algorithm was not used because Edgar (2004) demonstrated that MUSCLE was more accurate than other alignment methods. MEGA 6 also was used to find the best-fit model of nucleotide substitution (Posada and Crandall 2001). The model with the lowest Bayesian Information Criterion (BIC) scores is considered to best describe the substitution pattern of the data. The Kimura 2-parameter model with a gamma distribution was selected as the most appropriate model for the given data (BIC = 4362.019). Duplicate sequences were identified using the ElimDupes online program (http://hcv.lanl.gov/content/sequence/ELIMDUPES/elimdupes.html). Sequences that were 100% identical were removed in order to streamline phylogenetic tree construction. Two types of phylogenetic trees were constructed, and rooted, with a Chimney Swift (*Chaetura pelagica*) COI sequence (Table 1) available on GenBank. Reference COI sequences for Black-chinned and Ruby-throated Hummingbirds also were obtained from GenBank (Table 1).

Table 1: Ruby-throated Hummingbird and Black-chinned Hummingbird reference COI sequences for parental species available from GenBank; these samples were from outside the breeding range overlap. NCBI Accession Number is a unique identifier within GenBank. Haplotype ID groups the reference sequences with individuals from Appendix B according to exact sequence matches. Latitude and Longitude reference collection location. A *Chaetura pelagica* sequence available from GenBank was used to root the tree.

NCBI Accession Number	Date collected	Species ID	Latitude	Longitude	Haplotype ID
DQ433323	9/2/1994	*A. alexandri*	35.07º N	106.37º W	1
DQ432746	8/27/1993	*A. alexandri*	31.68º N	110.65º W	12
DQ433322	8/29/1994	*A. alexandri*	35.07º N	106.37º W	13
DQ433323	9/2/1994	*A. alexandri*	35.07º N	106.37º W	1
DQ433324	Unknown	*A. colubris*	43.33º N	80.15º W	9
DQ433325	5/13/2004	*A. colubris*	43.33º N	80.15º W	9
HM033219	8/18/2005	*A. colubris*	44.16º N	66.22º W	9
HM033220	8/26/2005	*A. colubris*	44.16º N	66.22º W	9
AY666220	Unknown	*A. colubris*	Unknown	Unknown	11
DQ433326	Unknown	*A. colubris*	43.33º N	80.15º W	14
HM033221	8/25/2005	*A. colubris*	44.16º N	66.22º W	15
EF373361	Unknown	*C. pelagica*	Unknown	Unknown	N/A

A Kimura 2-parameter model (Kimura 1980) was used to construct a Maximum Likelihood (ML) tree. The bootstrap consensus tree inferred from 1000 replicates was taken to represent the evolutionary history of the taxa analyzed (Felsenstein 1985). Branches corresponding to partitions reproduced in less than 50% of the bootstrap replicates were

collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) is shown next to the branches (Felsenstein 1985). Initial trees for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood approach, and then selecting the topology with a superior log likelihood value. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+G, parameter = 0.2121)). Seventy-two nucleotide sequences were analyzed, which included the *C. pelagica* tree root. Codon positions included were $1^{st} + 2^{nd} + 3^{rd}$ + Noncoding. There were 563 positions in the final dataset.

A Neighbor-Joining (NJ) tree (Saitou and Nei 1987) also was constructed. Bootstrapping was performed as described above for the ML tree. The evolutionary distances were computed using the Kimura 2-parameter method (Kimura 1980) and are in the units of the number of base substitutions per site. The rate variation among sites was modeled with a gamma distribution (shape parameter = 5). Seventy-two nucleotide sequences were analyzed, which included the *C. pelagica* tree root. Codon positions included were $1^{st} + 2^{nd} + 3^{rd}$ + Noncoding. All ambiguous positions were removed for each sequence pair. There were 563 positions in the final dataset.

Genetic Data Analysis

Maternal Species Identification Using Barcode Types

The barcode types for *Archilochus* identified in Tavares and Baker (2008) were used as a maternal species identification reference. Only five loci out of approximately 650 bp were fixed between species (Tavares and Baker 2008), thus the entire mitochondrial COI sequence was not used to identify maternal species. Maternal lineage was labeled as Ruby-throated Hummingbird

if the barcode type was CACAA, and was labeled as Black-chinned Hummingbird if the barcode type was TCTGG.

Per Locus and Sample Statistics for Microsatellite Loci

Number of alleles, and allele frequencies in each sample and overall were estimated in FSTAT v. 2.9.3 (Goudet 2001). Weighted (by sample size) and non-weighted frequencies were reported for overall allele frequencies. Gene diversity per locus and sample were estimated using an unbiased estimator (see Nei 1987 eq 7.39 p 164). Several individuals had missing data for one or two loci; statistics were calculated on the data set including all individuals and separately on a data set that did not include the individuals that were missing data.

Hardy-Weinberg per Locus and Sample for Microsatellite Loci

FSTAT also was used to estimate Hardy-Weinberg statistics. Alleles were randomized among individuals and within samples, and F_{IS} was used to compare the randomized data set to the observed data set. The nominal level for multiple tests of significance was set at 5% so that table-wide levels of significance could be observed.

Linkage Disequilibrium

In order to determine if loci were linked and needed to be combined, linkage disequilibrium was estimated using FSTAT. The statistic used to test the resulting linkage disequilibrium tables was the log-likelihood ratio G-statistic. Only individuals that were typed at both loci being tested were entered into the table. The nominal level for multiple tests was 5%.

Population Subdivision

Population structure analysis was performed using Structure v. 2.3.4 (Pritchard et al. 2000). Parental populations were not grouped for Structure analyses because the Black-chinned Hummingbird samples collected from west Texas to serve as a parental type did not show

sufficient microsatellite variation to distinguish them from the Ruby-throated Hummingbird parental samples (further explained in discussion). Structure results were analyzed with Structure Harvester (Earl and vonHoldt 2011) and graphs of the L(K) (mean likelihood of K (the number of actual populations)), rate of change of the likelihood distribution (L'(K)), absolute values in the rate of change (|L''(K)|), and change in K (ΔK) were estimated.

Conversion of Mitochondrial Single Nucleotide Polymorphisms to Allele Calls

Pritchard et al. (2003) discuss issues with using only linked sequence data such as mtDNA to perform population structure analysis. They suggest that the estimated population structure would likely represent the mtDNA tree, and the degree of uncertainty would likely be underestimated. However, if nuclear markers (such as microsatellites) were included, the mtDNA single nucleotide polymorphisms (SNPs) from each linked region could be recoded so that the linked loci would be represented as a single locus (Pritchard et al. 2003). Following this suggestion, SNPs in the COI sequence were isolated by copying and pasting them from MEGA into Microsoft Excel. Each nucleotide was assigned an arbitrary allele number (C = 100, A = 200, G = 300, T = 400). Each SNP was given a locus name (1 - 33) and all SNP loci were added to the microsatellite marker data set (36 loci total).

Allele frequencies, gene diversity, and linkage disequilibrium were estimated again for the new set of loci. Assumptions made by Structure were met because no significant linkage disequilibrium existed between loci (all p-values were ≥ 0.00005; p-value for 5% nominal level after Bonferroni corrections = 0.000046). Hardy-Weinberg calculations were not done on the new set because the circular mtDNA genome is haploid and violates the assumptions of Hardy-Weinberg equilibrium (Hardy 1908, Weinberg 1908). Population subdivision was estimated for the combined data set using Structure, and Structure Harvester was used to interpret data.

Inferred ancestry was used to identify population groupings, and a minimum estimated membership of 0.80 was used to assign individuals to a population. Individuals with estimated memberships between 0.2 and 0.8, or with overlapping probability intervals, were not assigned a population identification. Individuals in the combined microsatellite and mtDNA data set were labeled as hybrids if their field identification and structure identification were different. These hybrid locations were used to develop models to predict the extent of the hybrid zone.

GIS Methods

Current BioClim environmental conditions (Table 2) and elevation raster layers with a 2.5 arc-minute resolution were downloaded from WorldClim (Hijmans et al. 2005). Environmental layers were clipped to include only Canada, the United States, and Mexico. Duplicate hybrid locations were eliminated with the ENMTools v 1.3 (Warren et al. 2008, 2010). Any duplicate hybrid sites occurring within the same grid cell (2.5 arc-minute resolution) were eliminated from the data set in order to avoid spatial autocorrelation.

Hybrid locations and all environmental variables were imported into MaxEnt v 3.3.3k (Phillips et al. 2004, 2006) and habitat suitability was modeled with 10 replicates in order to determine which variables contributed the most to the model (York et al. 2014). Jackknife tests account for the dependencies between environmental variables by omitting each variable and refitting the model, as well as omitting all but one variable and refitting. The resulting training gain for each variable is then compared to the training gain when all variables are considered and is useful when determining which variables contribute most to the predictive power of the model.

Table 2: BioClim variables obtained from WorldClim. These represent a collection of global ecogeographical variables. The abbreviations for these variables are used in the remainder of the text.

Variable
BIO 1 = Annual Mean Temperature
BIO 2 = Mean Diurnal Range (Mean of monthly (max temp - min temp))
BIO 3 = Isothermality (BIO2/BIO7) (* 100)
BIO 4 = Temperature Seasonality (standard deviation *100)
BIO 5 = Max Temperature of Warmest Month
BIO 6 = Min Temperature of Coldest Month
BIO 7 = Temperature Annual Range (BIO5-BIO6)
BIO 8 = Mean Temperature of Wettest Quarter
BIO 9 = Mean Temperature of Driest Quarter
BIO 10 = Mean Temperature of Warmest Quarter
BIO 11 = Mean Temperature of Coldest Quarter
BIO 12 = Annual Precipitation
BIO 13 = Precipitation of Wettest Month
BIO 14 = Precipitation of Driest Month
BIO 15 = Precipitation Seasonality (Coefficient of Variation)
BIO 16 = Precipitation of Wettest Quarter
BIO 17 = Precipitation of Driest Quarter
BIO 18 = Precipitation of Warmest Quarter
BIO 19 = Precipitation of Coldest Quarter
Elevation

Training gain is similar to a measure of goodness of fit, and a high training gain is desirable for a good predictive model (greater deviation from random probability distribution). The eight variables with the highest mean training gain (for 10 replicates) when used in isolation had the most useful information by themselves and were retained for correlation analyses. The variable that decreased the mean training gain the most when it was omitted had the most information that was not present in the other variables and was also retained for correlation analyses.

The Pearson correlation coefficient per pair of environmental variables was calculated with ENMTools in order to eliminate correlated variables. A script was written in R v. 3.1.2 (R Development Core Team 2008) to automatically generate folders for all possible environmental variable combinations that had a correlation of < 0.5 so that they could be eliminated from the analysis. MaxEnt models were then generated for each remaining combination of environmental variables with raw outputs for model testing. Test points were a random sample selected from hybrid presence localities. The model selection function in ENMTools (Warren and Seifert 2011) was used to compare log likelihood, AIC (Akaike's Information Criterion) score, and AICc score (for small sample size) between the different environmental variable combination models. The environmental variables in the model with the lowest AICc score were used to generate the final MaxEnt model with ten-fold cross-validation. A habitat suitability map representing the point-wise mean of 10 output grids was reported and gives an estimate between 0 and 1 of probability of suitable habitat for hybrids. Hectares of suitable habitat (probability > 0.5) was calculated for each ecological niche model using QGIS.

GARP models were generated with the program openModeller Desktop v.1.1.0 (Muñoz et al. 2009). Hybrid localities were used for species presence points, and environmental variables

used to generate the suitability map were elevation and max temperature of warmest month (BIO 5, Table 2). Default settings were used on the GARP with best subsets model using the new openModeller implementation (Anderson et al. 2003). The habitat suitability map values were normalized (min = 0, max = 1) in order to be visually consistent with MaxEnt.

ENFA was run with the program BioMapper 4 (Hirzel et al. 2002b). Hybrid localities were projected onto a landscape and presence points were changed to 1 (absence points = 0). The resulting Boolean map was used as the species presence map, and environmental factors used were elevation and BIO 5. Covariance matrix, ecological niche factor analysis, and habitat suitability maps were all computed and k-fold cross-validation was performed (k = 2 using Huberty's rule). All other options were kept at default except for the habitat suitability algorithm, where geometric mean was used (Hirzel and Arlettaz 2003). The resulting habitat suitability map values were normalized (min = 0, max = 1) in order to be visually consistent with MaxEnt.

Raster outputs from each model were imported into QGIS and classified with the same color scale for visual consistency. In order to compare the accuracy of the habitat suitability maps between niche models, measures of model accuracy were calculated in R with the SDMTools package v. 1.1-221 (VanDerWal et al. 2014). Area under the curve (AUC), omission rate (as a proportion of true occurrences misidentified), sensitivity, specificity, proportion correctly identified (proportion of the presence and absence records correctly identified), and Kappa were calculated for a threshold value of 0.5.

Chapter III

RESULTS

Geographic distribution of plumages

One hundred and ninety-nine hummingbirds were captured at 30 sites (Figure 11). Sixty-one apparent Ruby-throated Hummingbirds were captured at one Alabama site, one Arkansas site, and seven Oklahoma sites. Nineteen adult males, one hatch year male, and 41 adult females were captured (Table 3). One hundred and thirty-seven apparent Black-chinned Hummingbirds were captured at 16 Texas sites and five Oklahoma sites. Twenty-two adult males, 24 hatch year males, 65 adult females, and 26 hatch year females were captured (Table 4). One adult male with hybrid plumage (L60384, Appendix A) was captured northeast of Decatur, TX (Figure 11). The only site where both species were observed was Chickasha, OK (Figure 11).

Table 3: Male and female Ruby-Throated Hummingbirds captured during this study.

	Hatch Year	After Hatch Year
Male	1	19
Female	0	41

Table 4: Male and female Black-chinned Hummingbirds captured during this study.

	Hatch Year	After Hatch Year
Male	24	22
Female	26	65

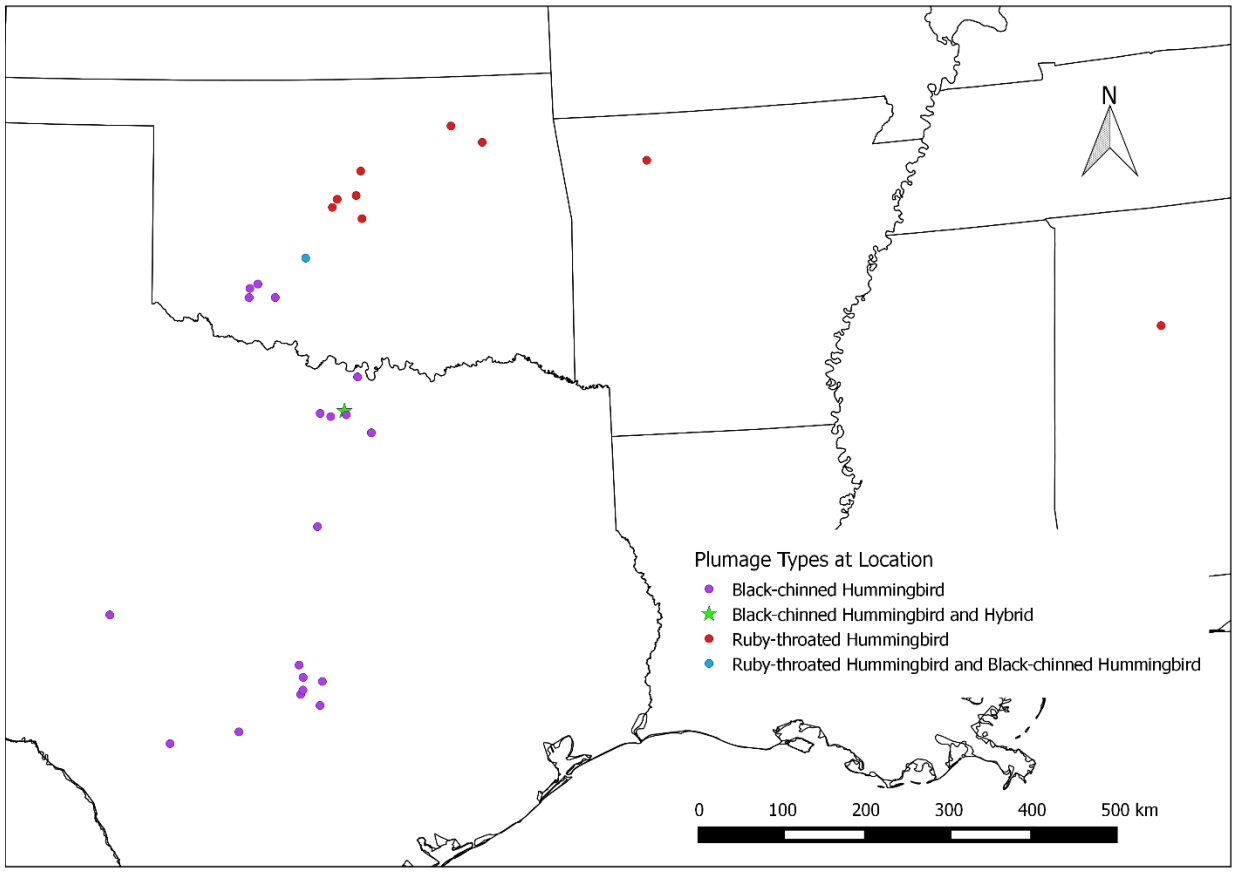

Figure 11: Locations of Black-chinned Hummingbirds and Ruby-throated Hummingbirds based on field identification of plumage. Black-chinned Hummingbird plumages were identified in southwest Oklahoma and Texas (purple). Ruby-throated Hummingbird plumages were identified in Oklahoma, Arkansas, and Alabama (red). Only one site had both species (blue). One hybrid plumage was identified at a Black-chinned Hummingbird site (green).

Sequencing

A 563 bp section of the cytochrome oxidase I (COI) gene was obtained for 187 of the 199 (94%) individuals. Low quantity or quality of DNA samples influenced the ability to obtain a full mitochondrial sequence in 12 individuals. After consolidation of duplicate sequences, 71 unique haplotypes remained (Table 5 and Appendix B). All duplicate haplotypes had the same barcode types as their haplotype representative. Haplotypes 1-11 had more than one member, whereas haplotypes 12-71 had only one member and many sequences that had ambiguous bases. Haplotype 5 was the most common being found in 47 individuals. Haplotype 9 was second most common being found in 37 individuals, and haplotype 11 was the third most common being found in 27 individuals. The apparent male hybrid individual had a unique haplotype (#58, Appendix B).

Haplotypes 2, 4, 5, 6, and 10 were not observed east of a North/South line from Guthrie (Logan County), OK to Austin (Travis County), TX (Figures 12 – 16). Haplotypes 9 and 11 occurred in Arkansas (Figure 15), and were the only two haplotypes identified in Alabama (Figure 16); however, haplotype 11 occurred as far west as Christoval (Tom Green County), TX (Figure 14) and haplotype 9 permeated sample sites in Oklahoma as far west as the Wichita Mountains (Figure 12). Haplotype 9 was not observed in Texas west of Muenster (Cooke County, Figures 14 and 15).

The three Black-chinned Hummingbird GenBank sequences that were downloaded to serve as reference samples fell into three haplotypes, two of which were unique (DQ432746 = 12, DQ433322 = 13). The Black-chinned Hummingbird GenBank sample (DQ433323) from Tijeras (Bernalillo County), NM was grouped into haplotype 1 with a field identified Black-chinned Hummingbird sample (E44516, Appendix A) from Chickasha (Grady County), OK.

Table 5: Unique haplotypes obtained from sequence data from 197 individuals (187 study samples and 10 GenBank samples). Only one representative of each haplotype is shown (Haplotype Representative) and identified with a number from 1 – 71 (Haplotype). Field ID contains the four letter alpha code for species (RTHU = Ruby-throated Hummingbird, BCHU = Black-chinned Hummingbird, and UNHY = Unknown Hybrid). Barcode type contains the nucleotide identification at the five bases used by Tavares and Baker (2008) for species identification and includes mixed base identifications (IUPAC Ambiguity Codes: Y = A or G, M = G or T, R = C or T). Quantity is the number of individuals with each haplotype.

Haplotype Representative	Haplotype	Field ID	Barcode Type	Quantity
E44516	1	BCHU	TCTGG	2
H25371	2	BCHU	CCTAA	3
L60389	3	BCHU	CACAA	4
H25373	4	BCHU	CCTAA	2
H25244	5	RTHU	CCTAA	47
L4100-68541	6	BCHU	CCTAA	2
H25396	7	RTHU	CACAA	2
L34922	8	BCHU	CACAA	9
3100H-35801	9	RTHU	CACAA	37
H25360	10	BCHU	CCTAA	2
4100L-68624	11	RTHU	CACAA	27
GB Accession DQ432746	12	BCHU	TCTGG	1
GB Accession DQ433322	13	BCHU	TCTGG	1
GB Accession DQ433326	14	RTHU	CACAA	1
GB Accession HM033221	15	RTHU	CACAA	1
H25245	16	RTHU	CACAA	1
H25246	17	RTHU	CACAA	1
H25248	18	RTHU	CMYAA	1
H25249	19	BCHU	CATAA	1
H25250	20	BCHU	CMYAA	1
H25256	21	BCHU	CAYAA	1

Table 5: Continued (IUPAC Ambiguity Codes: Y = A or G, M = G or T, R = C or T).

Haplotype Representative	Haplotype	Field ID	Barcode Type	Quantity
H25259	22	BCHU	CAYAA	1
H25260	23	BCHU	CACAA	1
H25330	24	BCHU	YMYAA	1
H25331	25	BCHU	YCTAA	1
H25334	26	BCHU	TCYAA	1
H25335	27	RTHU	CAYAA	1
H25336	28	BCHU	TCTAA	1
H25337	29	BCHU	TCTGG	1
H25338	30	BCHU	CACAA	1
H25339	31	BCHU	YMYRR	1
H25341	32	BCHU	TATRR	1
H25342	33	BCHU	CCYRR	1
H25348	34	BCHU	CACAA	1
H25358	35	BCHU	CMYAA	1
H25362	36	BCHU	CCTAA	1
H25369	37	BCHU	CMYAA	1
H25370	38	BCHU	CCTAA	1
H25372	39	BCHU	CCTAA	1
H25376	40	BCHU	YCTRR	1
H25382	41	BCHU	CMYAA	1
H25383	42	BCHU	CMYAA	1
H25387	43	BCHU	CCTAA	1
H25391	44	BCHU	CMYAA	1
H25393	45	BCHU	YMYAA	1
H25398	46	RTHU	CMYAA	1
H25400	47	RTHU	CACAA	1
L34341	48	BCHU	CMYAA	1
L34342	49	BCHU	CMYAA	1

Table 5: Continued (IUPAC Ambiguity Codes: Y = A or G, M = G or T, R = C or T).

Haplotype Representative	Haplotype	Field ID	Barcode Type	Quantity
L34403	50	BCHU	CCTAA	1
L34410	51	BCHU	CCTAA	1
L34414	52	BCHU	CMYAA	1
L34910	53	BCHU	CMYAA	1
L35654	54	BCHU	YMYRR	1
L4100-68539	55	BCHU	CCTAA	1
L4100-68551	56	BCHU	CACAA	1
L60383	57	BCHU	CACAA	1
L60384	58	UNHY	CACAA	1
L60388	59	BCHU	CACAA	1
L60393	60	BCHU	CCTAA	1
L60395	61	BCHU	YMYRR	1
L60399	62	BCHU	CMYAA	1
L60401	63	BCHU	CACAA	1
L60404	64	BCHU	CACAA	1
L60409	65	BCHU	CCTAA	1
L60412	66	BCHU	CMYAA	1
L60454	67	RTHU	CACAA	1
L60457	68	RTHU	CACAA	1
L60463	69	RTHU	CCCAA	1
L60474	70	BCHU	CACAA	1
L60475	71	BCHU	CACAA	1

Four of the Ruby-throated Hummingbird GenBank sequences (DQ433324 - DQ433325, HM033219 - HM033220) that were downloaded to serve as reference samples grouped with the Ruby-throated Hummingbird parental samples from Alabama in haplotype 9. Ruby-throated Hummingbird GenBank sample AY666220 grouped with haplotype 11, and DQ433326 and HM033221 were unique haplotypes (14 and 15, respectively).

Base calls for the barcode types published by Tavares and Baker (2008) were CACAA for Ruby-throated Hummingbirds and TCTGG for Black-chinned Hummingbirds. Ninety out of 187 (48%) individuals sequenced had the Ruby-throated Hummingbird barcode type, whereas only two (1%) had the Black-chinned Hummingbird barcode type (Table 5). Twenty-seven (14%) individuals had ambiguous bases at one or more of the barcode loci that are represented by IUPAC (International Union of a Pure and Applied Chemistry) codes. The remaining 68 (36%) sequenced individuals had barcode types not identified by Tavares and Baker (2008) and included CCTAA (65 individuals), CCCAA (1), CATAA (1), and TCTAA (1).

Thirty-three single nucleotide polymorphisms (SNPs) in the mtDNA sequences from all individuals were identified, including the five fixed loci described by Tavares and Baker (2008). The remaining 28 loci were not fixed for either group of parental species representatives. Ninety individuals had the CACAA barcode type, indicative of Ruby-throated Hummingbirds, as described by Tavares and Baker (2008). Fifty-one of the individuals with the CACAA barcode type had Ruby-throated Hummingbird plumages, 38 had Black-chinned Hummingbird plumages, and one had a hybrid plumage. Two samples (H25337 and E44516) from southwest OK (Grady Co., Comanche Co.) were the only individuals that had the same barcode type (TCTGG) as the Black-chinned Hummingbirds in Tavares and Baker (2008), and both had Black-chinned

Hummingbird plumages. The most prevalent of the intermediate barcode types (CCTAA) consisted almost entirely of Black-chinned Hummingbird phenotypes.

Haplotypes 3, 7, 8, 9, 11, 23, 30, 34, 47, 56-59, 60-63, 67-68, and 70-71 were associated with the published Ruby-throated Hummingbird barcode type (CACAA). Haplotypes 1 and 29 were associated with the published Black-chinned Hummingbird barcode type (TCTGG). The individuals with the previously unidentified barcode types (CCTAA) were members of haplotypes 2, 4, 5, 6, 10, 36, 38, 39, 43, 50, 51, 55, 60, and 65. Haplotypes 19, 28, and 69 represent the previously unidentified barcode types CATAA, TCTAA, and CCCAA, respectively. The individual with a magenta gorget (L60384) had a Ruby-throated Hummingbird (CACAA) barcode type. However, the method of hybrid identification used in this study is reliant upon contrasting barcode type and field identification.

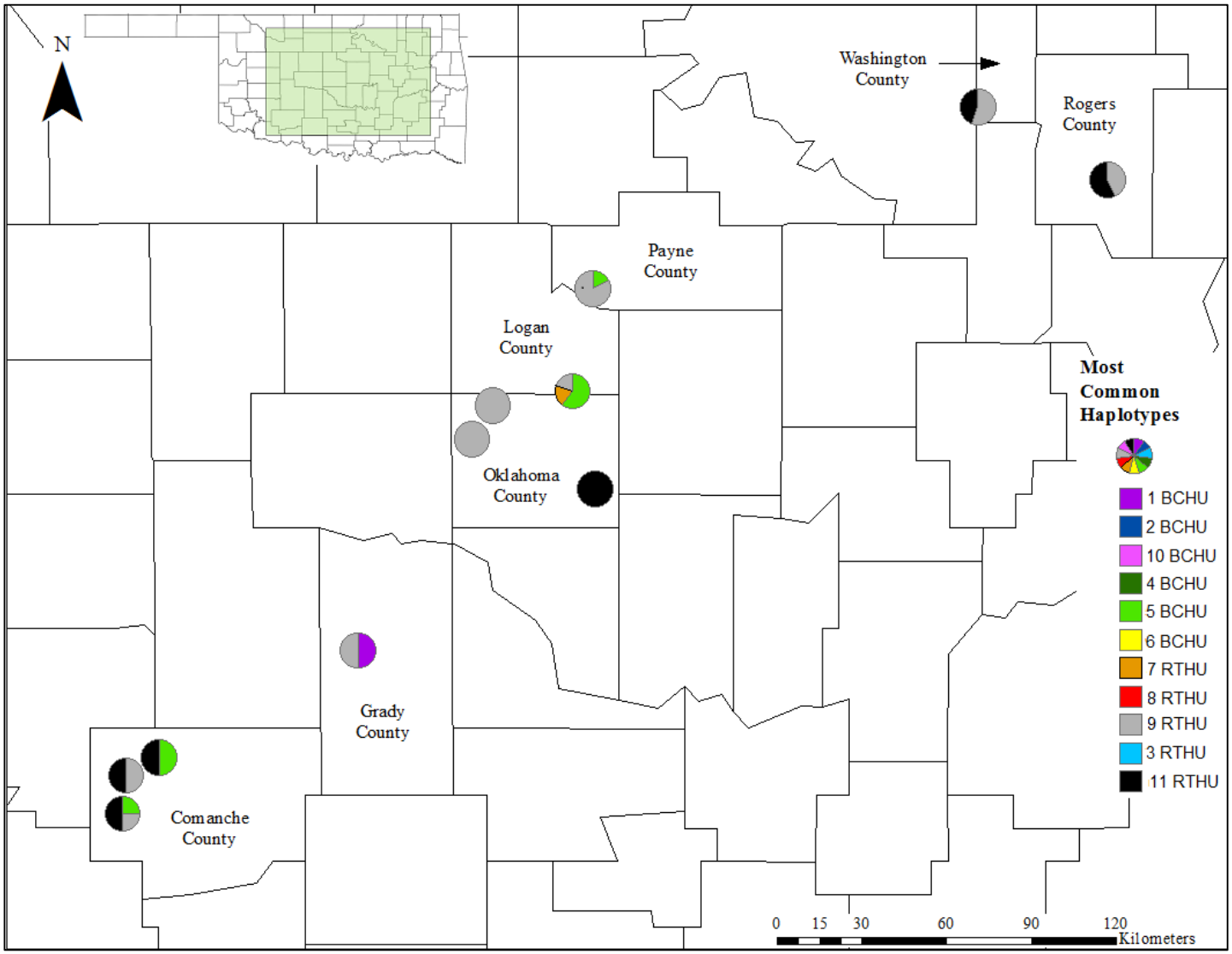

Figure 12: Haplotype map of samples collected in Oklahoma. Haplotype numbers correspond to Table 5; four letter alpha code represents barcode type that corresponds to each haplotype. Only the most common haplotypes were utilized in the haplotype maps. See Appendix B for haplotypes present in all individuals.

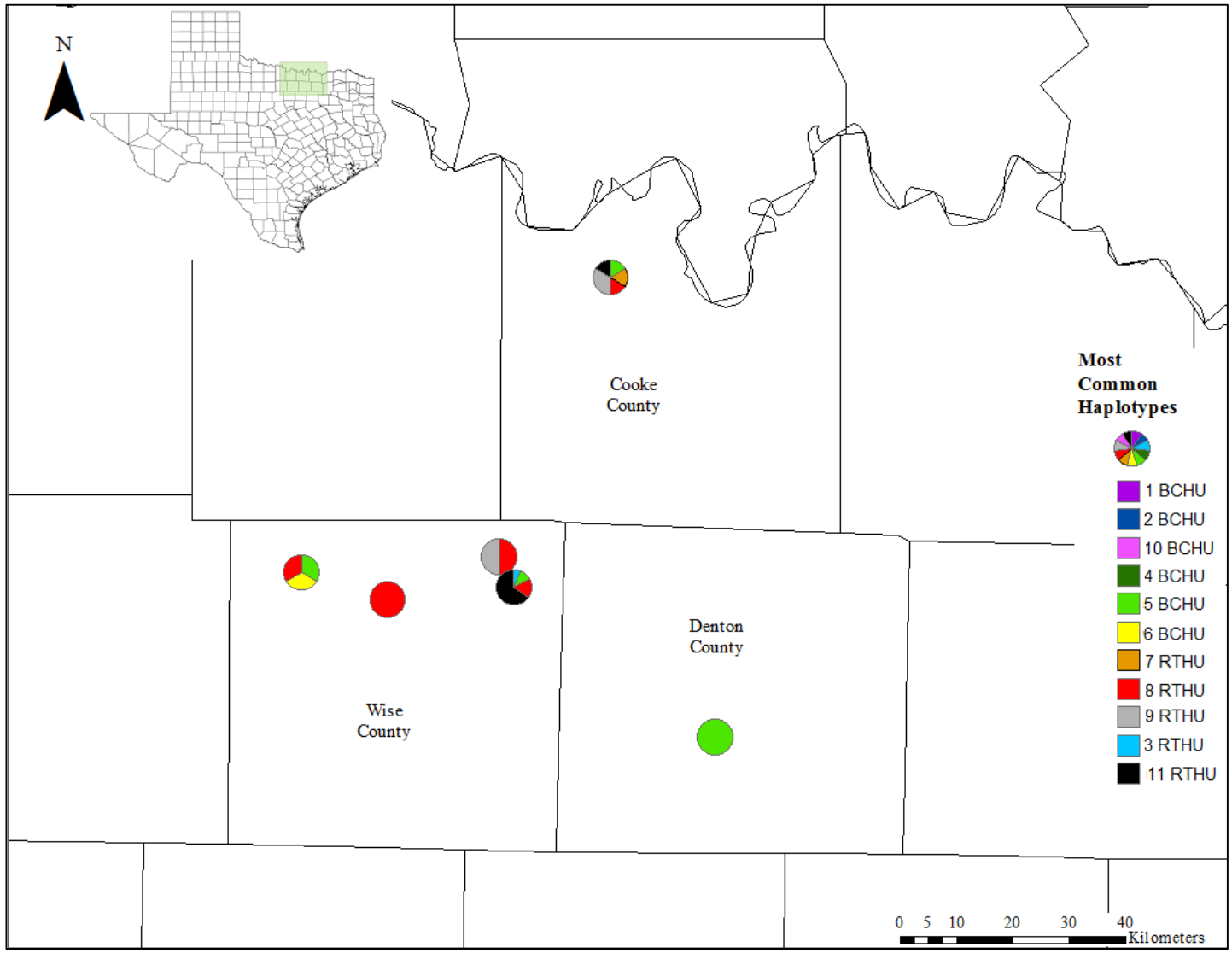

Figure 13: Haplotype map of samples collected in north central Texas. Haplotype numbers correspond to Table 5; four letter alpha code represents barcode type that corresponds to each haplotype. Only the most common haplotypes were utilized in the haplotype maps. See Appendix B for haplotypes present in all individuals.

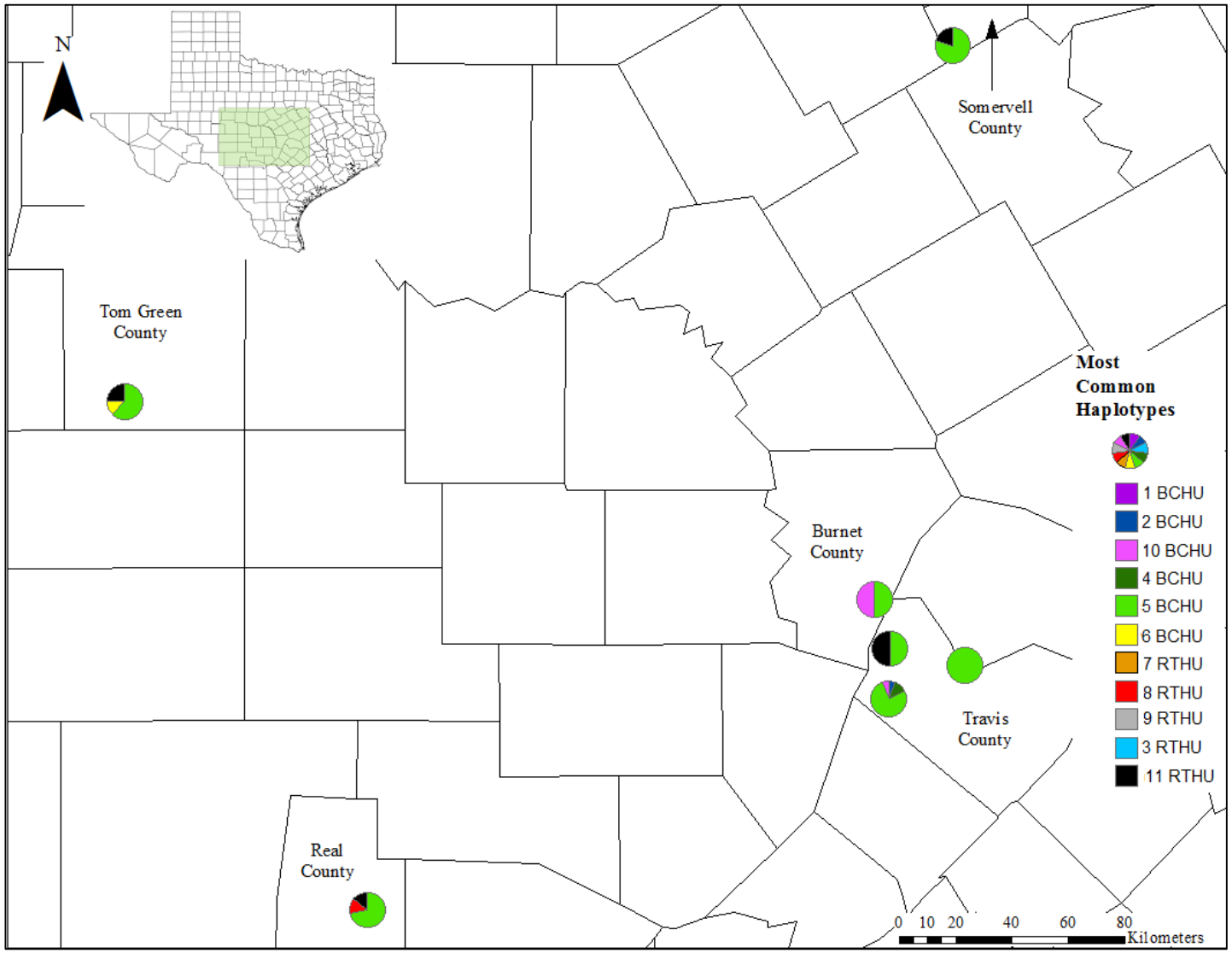

Figure 14: Haplotype map of samples collected in south central and west Texas. Haplotype numbers correspond to Table 5; four letter alpha code represents barcode type that corresponds to each haplotype. Only the most common haplotypes were utilized in the haplotype maps. See Appendix B for haplotypes present in all individuals.

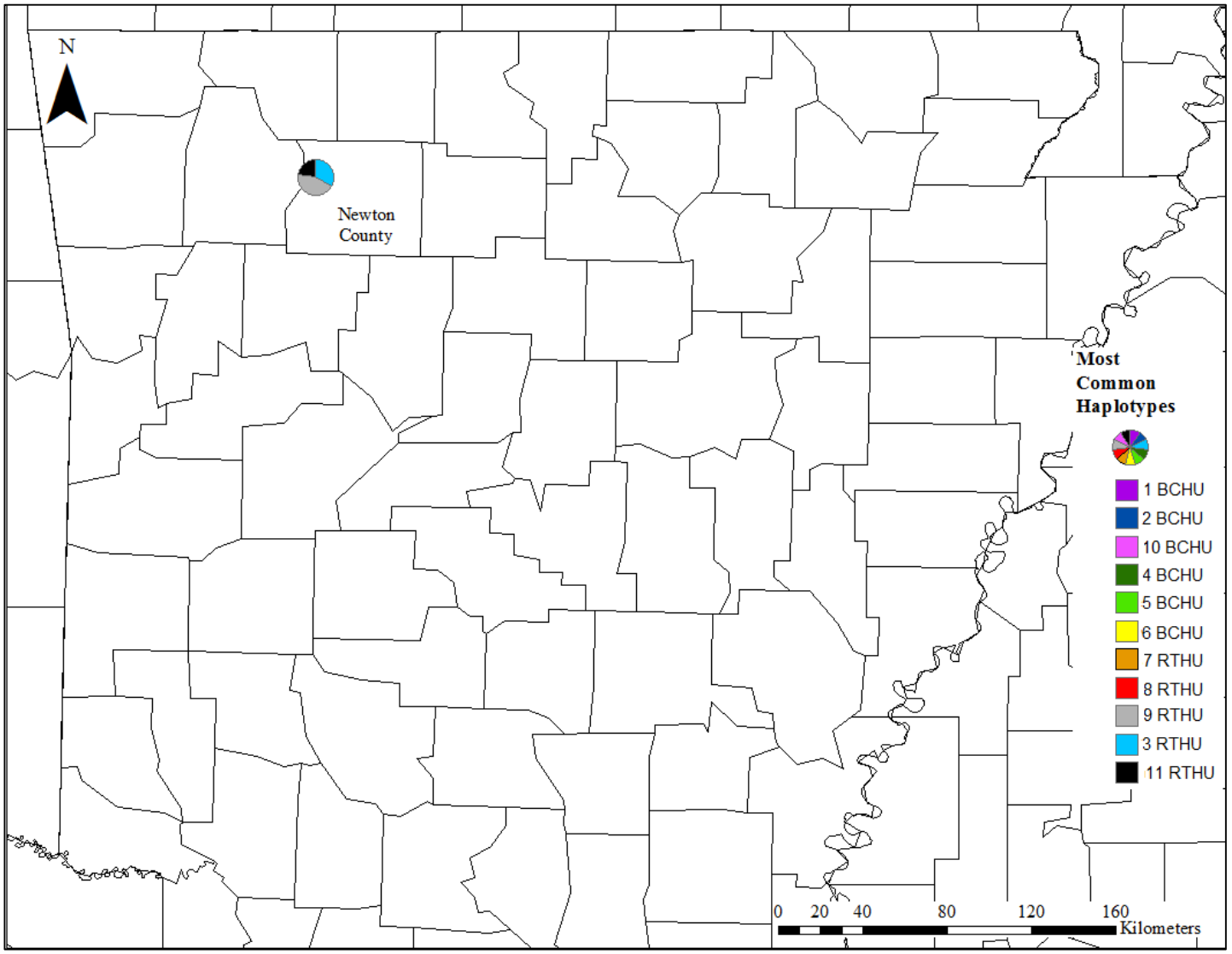

Figure 15: Haplotype map of samples collected in Arkansas. Haplotype numbers correspond to Table 5; four letter alpha code represents barcode type that corresponds to each haplotype. Only the most common haplotypes were utilized in the haplotype maps. See Appendix B for haplotypes present in all individuals.

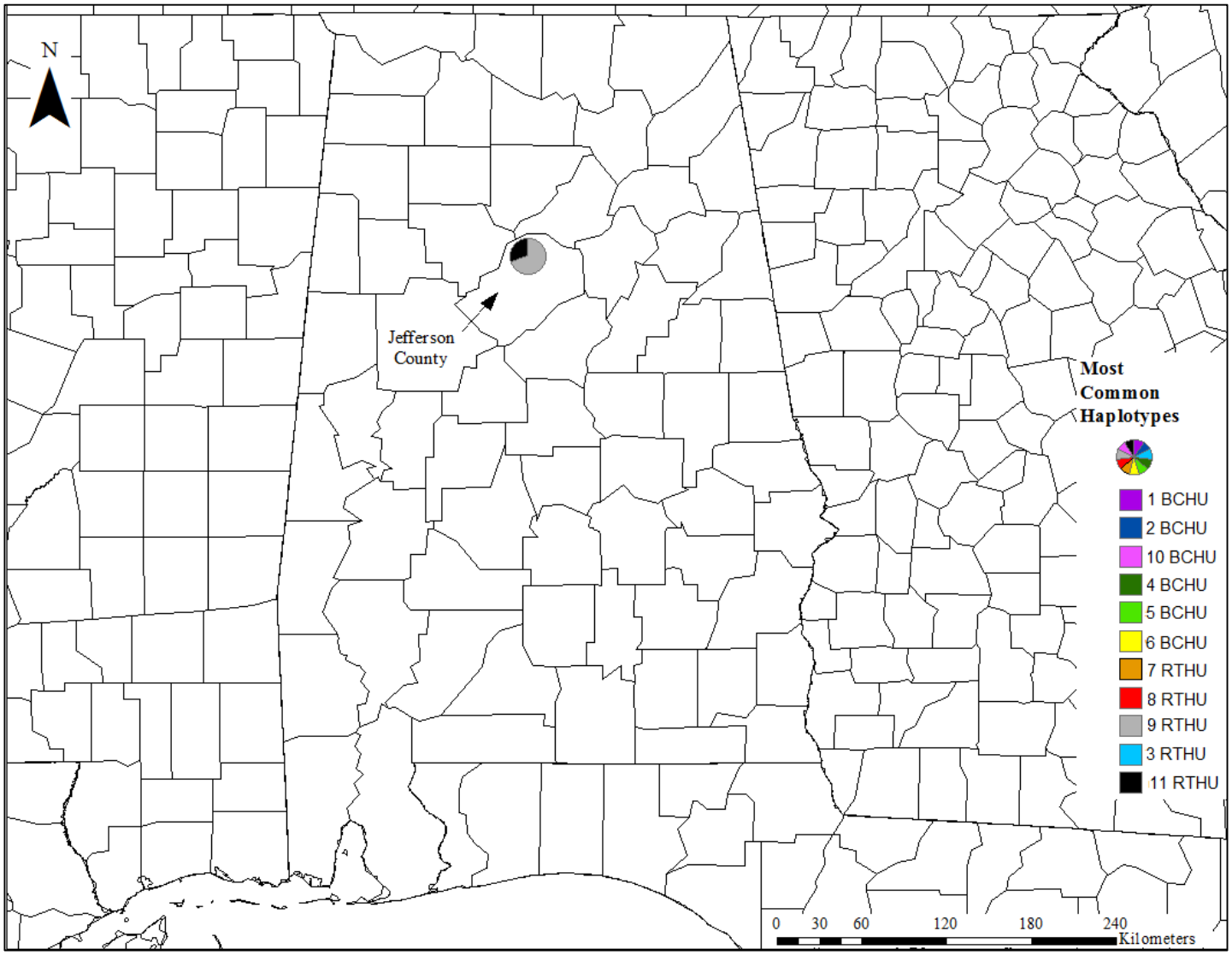

Figure 16: Haplotype map of samples collected in Alabama. Haplotype numbers correspond to Table 5; four letter alpha code represents barcode type that corresponds to each haplotype. Only the most common haplotypes were utilized in the haplotype maps. See Appendix B for haplotypes present in all individuals.

Phylogenetic Tree Construction

Black-chinned Hummingbird GenBank samples (haplotypes 1, 12, and 13) clustered together in the NJ tree (Figure 18). The ML tree grouped most of the Ruby-throated Hummingbird GenBank samples (haplotypes 9, 11, and 15) in the same cluster, but one sample (haplotype 14) was separated from other Ruby-throated Hummingbird haplotypes by Black-chinned Hummingbird haplotypes (bootstrap support < 50). The NJ tree grouped six Ruby-throated Hummingbird GenBank samples (haplotypes 9, 14, and 15) in the same cluster (bootstrap support < 50), but the remaining sample (haplotype 11) occurred in a separate cluster.

On the NJ tree (Figure 18), most of the haplotypes outside of the Black-chinned Hummingbird cluster were associated with the CACAA Ruby-throated Hummingbird barcode type. Haplotypes with ambiguous bases at one or more of the barcode sites had significantly more undefined sites in the remaining COI sequence and should be interpreted with caution. Four haplotypes at the bottom of the blue cluster represent the TCTGG Black-chinned Hummingbird barcode type. All remaining haplotypes that occurred in the blue cluster were intermediate barcode types, ambiguous at the relevant sites, or Ruby-throated Hummingbird barcode types. Two Ruby-throated Hummingbird CACAA barcode types were present in this cluster (haplotypes 7 and 47). Haplotypes 2, 4, 5, 6, and 10, which appear to be restricted to western *Archilochus* populations, grouped together at the top of the blue cluster and represented individuals with the CCTAA intermediate type. Clustering of the CCTAA intermediate barcode type with the TCTGG Black-chinned Hummingbird barcode type suggests that the CCTAA barcode type is more closely related to Black-chinned Hummingbirds than Ruby-throated Hummingbirds, although support is low.

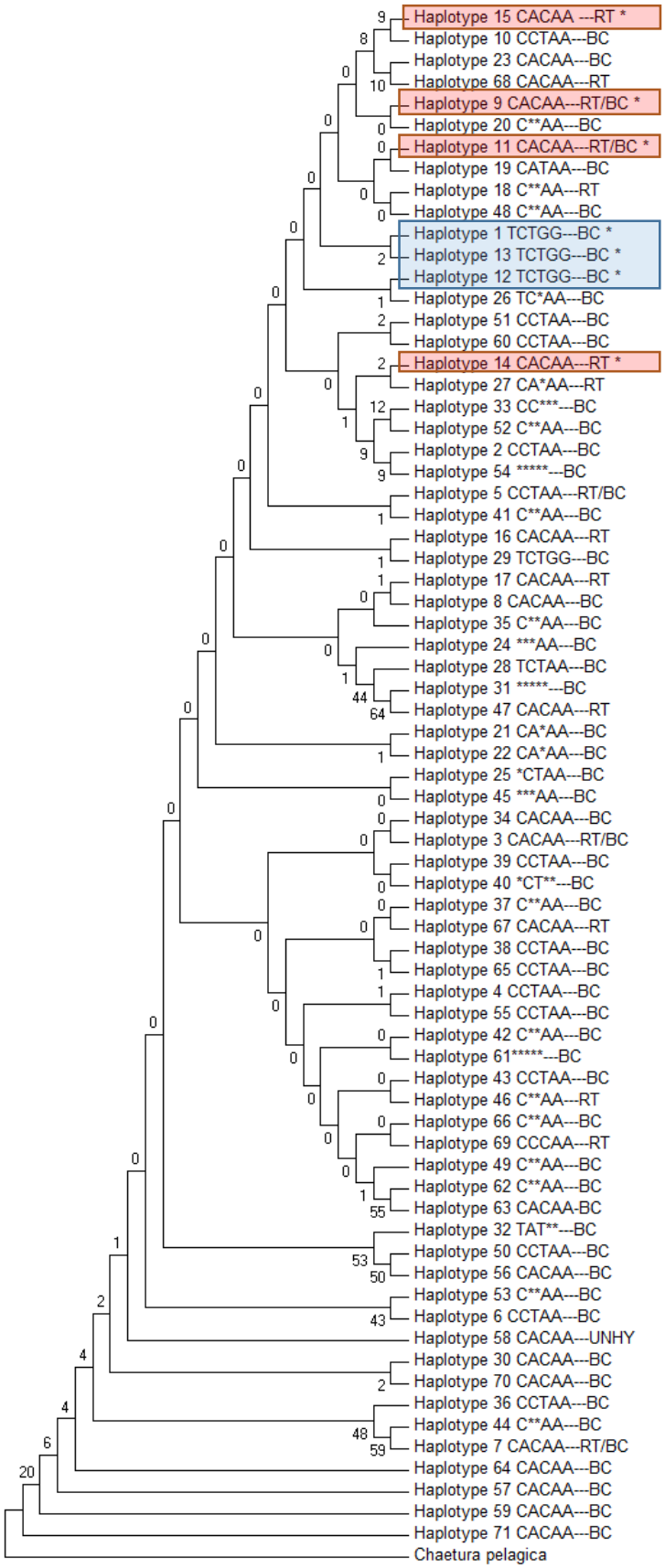

Figure 17: Maximum Likelihood tree constructed using MEGA v. 6 (Tamura et al. 2013). Blue denotes Black-chinned Hummingbird GenBank haplotypes, and pink denotes Ruby-throated Hummingbird GenBank haplotypes. Haplotype #, barcode type associated with haplotype (* denotes ambiguous base), and plumages associated with haplotype (RT represents Ruby-throated, BC represents Black-chinned, RT/BC means both plumages occurred in haplotype, and UNHY is the male with hybrid plumage) are listed at the end of each branch. Haplotype numbers correspond to Table 5. Haplotypes that contained GenBank reference samples are denoted with an asterisk after the plumage type. Bootstrap values are shown at each node. *Chaetura. pelagica* serves as the tree root.

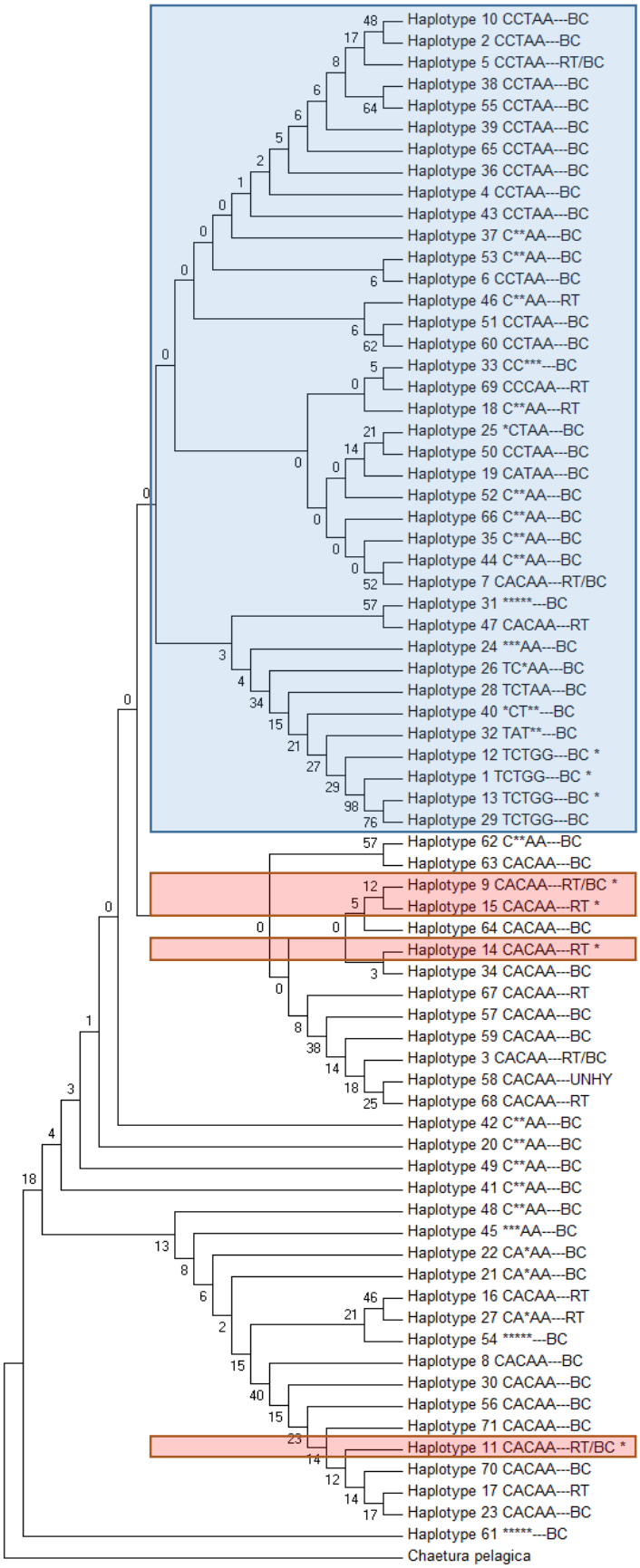

Figure 18: Neighbor-joining tree constructed using MEGA v. 6 (Tamura et al. 2013). Blue denotes Black-chinned Hummingbird cluster, and pink denotes Ruby-throated Hummingbird GenBank samples. Haplotype #, barcode type associated with haplotype (* denotes ambiguous base), and plumages associated with haplotype (RT represents Ruby-throated, BC represents Black-chinned, RT/BC means both plumages occurred in haplotype, and UNHY is the male with hybrid plumage) are listed at the end of each branch. Haplotype numbers correspond to Table 5. Haplotypes that contained GenBank reference samples are denoted with an asterisk after the plumage distinctions. Bootstrap values are shown at each node. *Chaetura. pelagica* serves as tree root.

Population Structure Analysis

Per Locus and Sample Statistics

Allele frequency and gene diversity analyses revealed two loci (11 and 14) with a fairly equal distribution of alleles (frequencies between 0.4 and 0.6) and 27 loci with a high prevalence (frequency > 0.90) of one allele (Table 6). No genotypes were identified that were representative of either parent species. No significant linkage disequilibrium existed between the three microsatellite markers and the 33 single nucleotide polymorphisms (p-values were ≥ 0.000046).

Population Subdivision

Population structure analysis of the combined mtDNA SNPs and microsatellites suggested two clusters (Figure 19). Structure analysis using the combined dataset assigned the Alabama Ruby-throated Hummingbird individuals to the same cluster (Appendix B), which the analysis with only three microsatellite loci failed to do. Black-chinned Hummingbird types in Christoval (Tom Green County), TX were not all grouped into the same cluster, although seven out of ten individuals were clustered together (Appendix B). Three individuals at that location were assigned to a cluster containing the Ruby-throated Hummingbird parental types. This suggests that we did not sample far enough outside of the hybrid zone to observe a pure parental Black-chinned Hummingbird population.

Structure analysis grouped 104 individuals with all of the parental Ruby-throated Hummingbird samples from Alabama, suggesting that this cluster represents Ruby-throated Hummingbirds (cluster one). Structure assigned 65 individuals to a separate cluster that included seven of the parental Black-chinned Hummingbird representatives from west Texas (cluster two). However, three of the Black-chinned Hummingbird parental representatives clustered with the Ruby-throated Hummingbirds. The members of cluster two possessed the most common

Table 6: Allele frequencies and gene diversity for three microsatellite loci and 33 single nucleotide polymorphisms (SNPs) within the cytochrome oxidase I gene were calculated using FSTAT v. 2.9.3 (Goudet 2001). Marker represents either the name of the three microsatellite loci (Tro 2, 4, 10) assigned by Lance et al. (2008) or the name of the SNP identified in this study (Locus 1 – 33). Number of Samples refers to the number of collected DNA samples that have data for the given loci (out of 187). Allele Size for the three microsatellite markers represents the size of the allele in base pairs, whereas it represents a specific nucleotide for the SNP loci (100 = C, 200 = A, 300 = G, 400 = T). An asterisk represents high prevalence of a single allele in the population.

Marker	Number of Samples	Allele Size	Frequency	Gene Diversity
Tro 2	179	p: 163	0.014	0.17
		p: 164	0.078	
		p: 166	0.908 *	
Tro 4	178	p: 169	0.037	0.076
		p: 172	0.003	
		p: 175	0.961 *	
Tro 10	176	p: 225	0.009	0.35
		p: 226	0.011	
		p: 228	0.069	
		p: 229	0.003	
		p: 230	0.003	
		p: 232	0.797	
		p: 236	0.106	
		p: 240	0.003	
Locus 1	187	p: 100	0.995 *	0.011
		p: 400	0.005	
Locus 2	187	p: 200	0.995 *	0.011
		p: 300	0.005	
Locus 3	187	p: 200	0.995 *	0.011
		p: 300	0.005	

Table 6: Continued

Marker	Number of Samples	Allele Size	Frequency	Gene Diversity
Locus 4	177	p: 200	0.633	0.467
		p: 300	0.367	
Locus 5	165	p: 100	0.994 *	0.012
		p: 200	0.006	
Locus 6	182	p: 200	0.269	0.396
		p: 300	0.731	
Locus 7	187	p: 100	0.005	0.011
		p: 400	0.995 *	
Locus 8	177	p: 100	0.633	0.467
		p: 400	0.367	
Locus 9	187	p: 200	0.995 *	0.011
		p: 300	0.005	
Locus 10	180	p: 100	0.972 *	0.054
		p: 400	0.028	
Locus 11	168	p: 100	0.435	0.494
		p: 200	0.565	
Locus 12	185	p: 100	0.011	0.022
		p: 400	0.989 *	
Locus 13	186	p: 100	0.011	0.021
		p: 400	0.989 *	
Locus 14	163	p: 100	0.558	0.496
		p: 400	0.442	
Locus 15	169	p: 200	0.385	0.476
		p: 300	0.615	
Locus 16	185	p: 100	0.995 *	0.011
		p: 400	0.005	
Locus 17	187	p: 100	0.995 *	0.011
		p: 400	0.005	

Table 6: Continued

Marker	Number of Samples	Allele Size	Frequency	Gene Diversity
Locus 18	172	p: 200	0.256	0.383
		p: 300	0.744	
Locus 19	187	p: 100	0.005	0.011
		p: 400	0.995 *	
Locus 20	169	p: 200	0.367	0.467
		p: 300	0.633	
Locus 21	187	p: 100	0.005	0.011
		p: 400	0.995 *	
Locus 22	181	p: 200	0.989 *	0.022
		p: 300	0.011	
Locus 23	181	p: 200	0.989 *	0.022
		p: 300	0.011	
Locus 24	187	p: 100	0.989 *	0.021
		p: 400	0.011	
Locus 25	186	p: 200	0.995 *	0.011
		p: 300	0.005	
Locus 26	187	p: 100	0.984 *	0.032
		p: 400	0.016	
Locus 27	180	p: 200	0.017	0.033
		p: 400	0.983 *	
Locus 28	184	p: 100	0.011	0.022
		p: 400	0.989 *	
Locus 29	186	p: 100	0.989 *	0.021
		p: 400	0.011	
Locus 30	187	p: 200	0.005	0.011
		p: 300	0.995 *	
Locus 31	187	p: 100	0.995 *	0.011
		p: 400	0.005	

Table 6: Continued

Marker	Number of Samples	Allele Size	Frequency	Gene Diversity
Locus 32	178	p: 100	0.011	0.022
		p: 400	0.989 *	
Locus 33	186	p: 100	0.989 *	0.021
		p: 400	0.011	

intermediate mtDNA type (CCTAA), which contained mostly Black-chinned Hummingbird plumages. The haplotypes of the 65 individuals in cluster two were geographically limited to the published Black-chinned Hummingbird breeding range as well. Given the prevalence of the Black-chinned Hummingbird plumage in cluster two, the geographic distribution of the haplotypes in cluster two, and that many of the Black-chinned Hummingbird parental representatives were assigned to cluster two, it was assumed that cluster two represented Black-chinned Hummingbirds. The remaining 18 individuals had mixed proportions of clusters or overlapping confidence intervals and were labeled as "uncertain." The two Black-chinned Hummingbird (TCTGG) types identified in Grady and Comanche counties in Oklahoma fell into this "uncertain" group.

Fifty-three individuals had Ruby-throated Hummingbird plumage and genetic types, and 62 individuals had Black-chinned Hummingbird plumage and genetic types. Fifty-four individuals had conflicting plumage types and genetic types (N = 4 Ruby-throated Hummingbird plumages and Black-chinned Hummingbird genetic IDs, N = 50 Black-chinned Hummingbird plumages and Ruby-throated Hummingbird genetic IDs) and were used as hybrid presence points in the ecological niche modeling comparisons. Individuals labeled as "uncertain" were not used in ecological niche modelling because their genetic identification was questionable and could not be compared to the plumage type for hybrid identification.

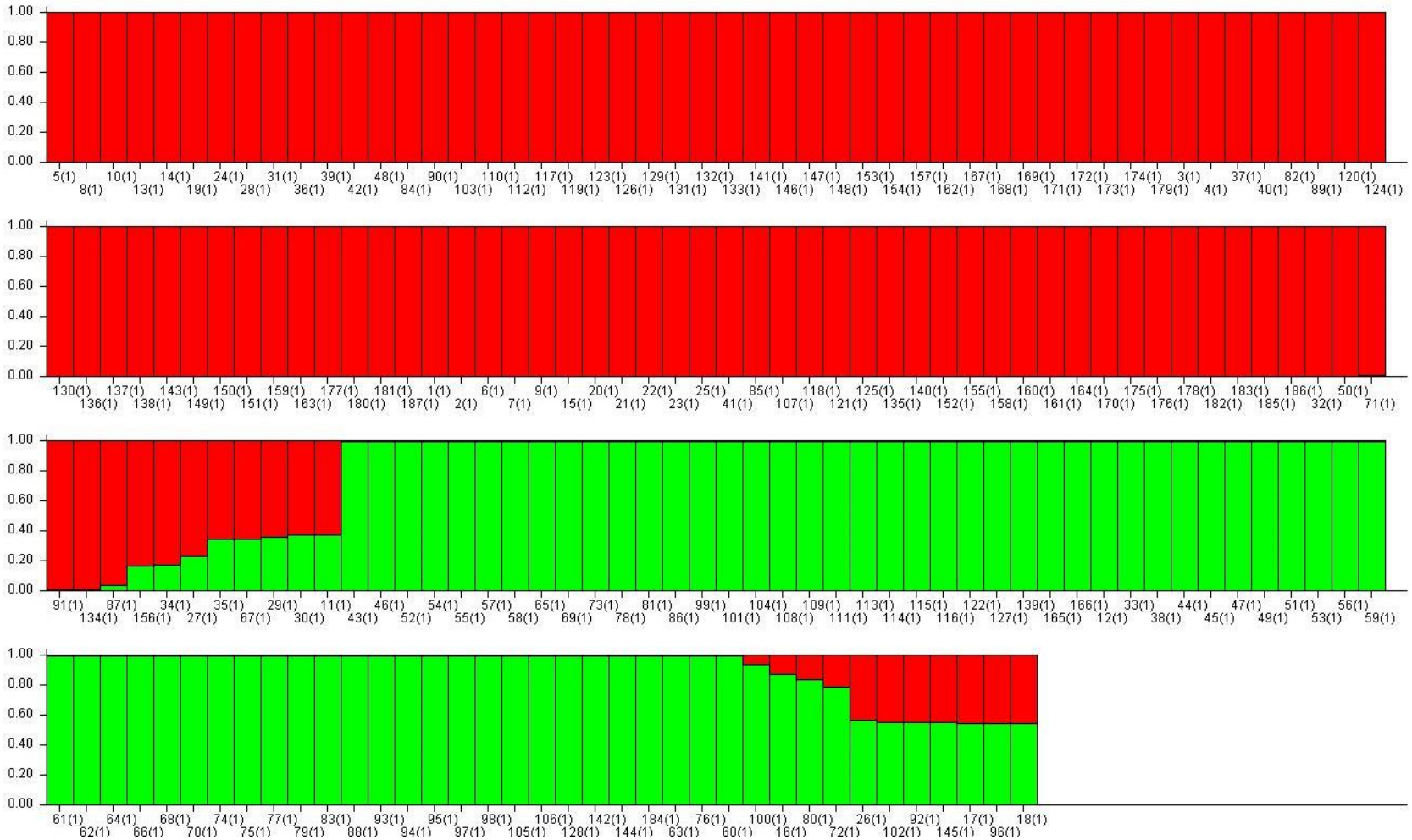

Figure 19: Summary plot of Q values estimated using Structure v. 2.3.4 (Pritchard et al. 2000). Each individual is represented by a single vertical bar (number under bar corresponds to Appendix B) broken into colored segments that are proportional to each of the inferred clusters. Red represents the Ruby-throated Hummingbird cluster, and green represents the Black-chinned Hummingbird cluster. Individuals with Q values between 0.2 and 0.8 were considered to have uncertain genetic identification and not used in ecological niche modelling.

Ecological Niche Modeling

Fifty-four hybrids were identified at 18 sites, but after elimination of redundant hybrid locations within a 2.5 arc-minute square, only 17 presence points remained (Figure 20). The eight environmental variables with the highest mean training gain (when only that variable was considered) were BIO 1, 3, 5, 6, 8, 9, 10, and 11 (Figure 21). The variable that lowered the training gain the most when removed from the analysis was elevation (Figure 22). The combination of uncorrelated environmental variables (correlation coefficient < 0.5) that produced the model with the lowest AICc (390.814, Table 7) was elevation and max temperature of warmest month (BIO 5).

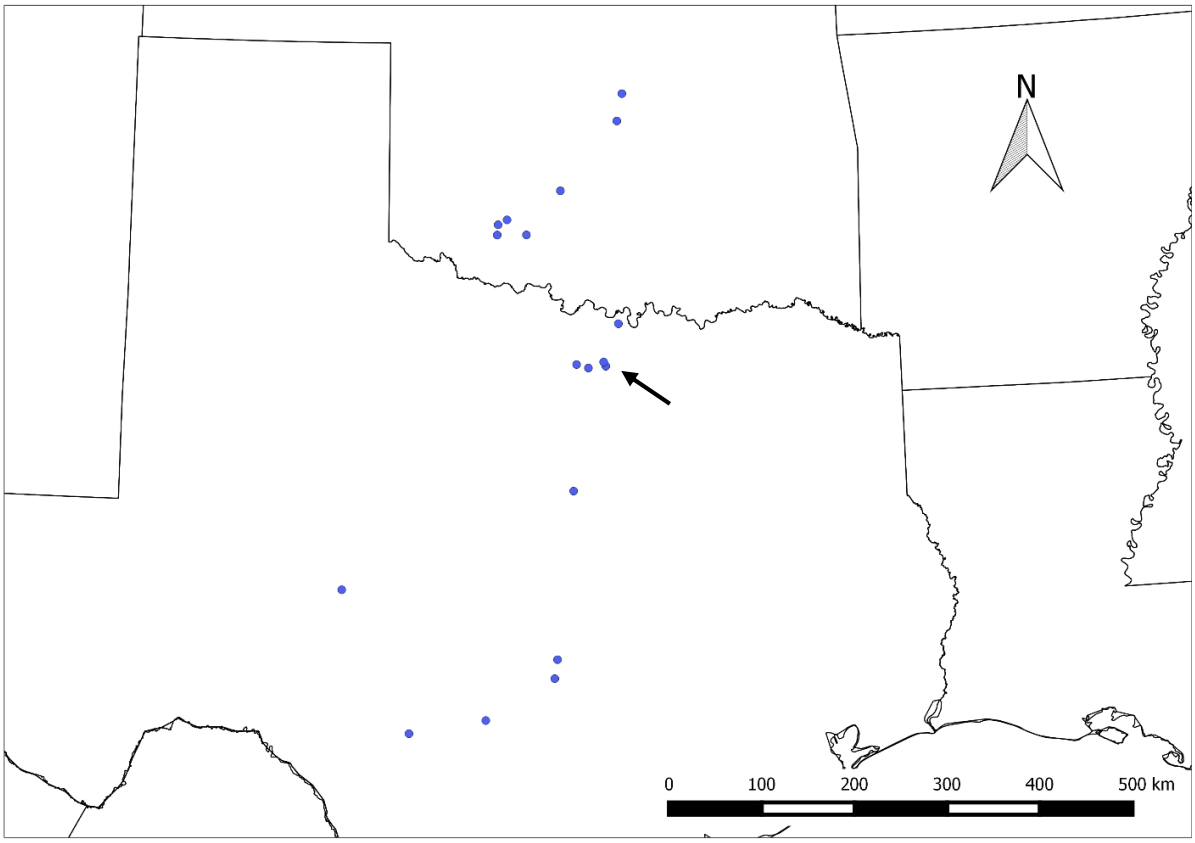

Figure 20: Hybrid locations. Fifty-four hybrids were identified at 18 sites across Oklahoma and Texas. One site was omitted due to close proximity of another site (indicated by arrow).

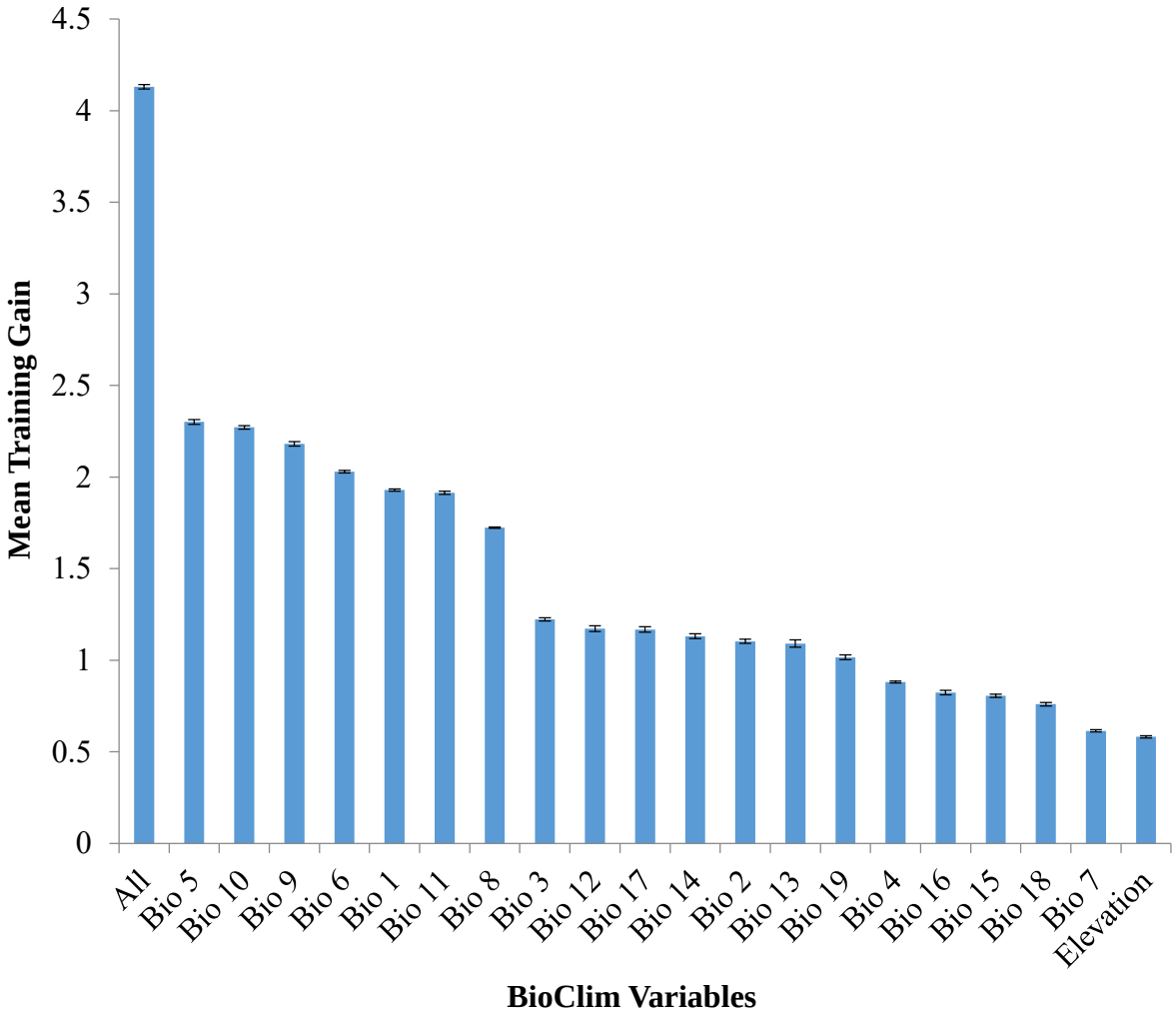

Figure 21: Jackknife test of variable importance with mean training gain (over 10 runs) for all variables and for each variable in isolation. The eight variables with the highest mean training gain when used in isolation appear to have the most useful information by themselves (Bio 1, 3, 5, 6, 8, 9, 10, and 11). BioClim variables correspond to Table 2.

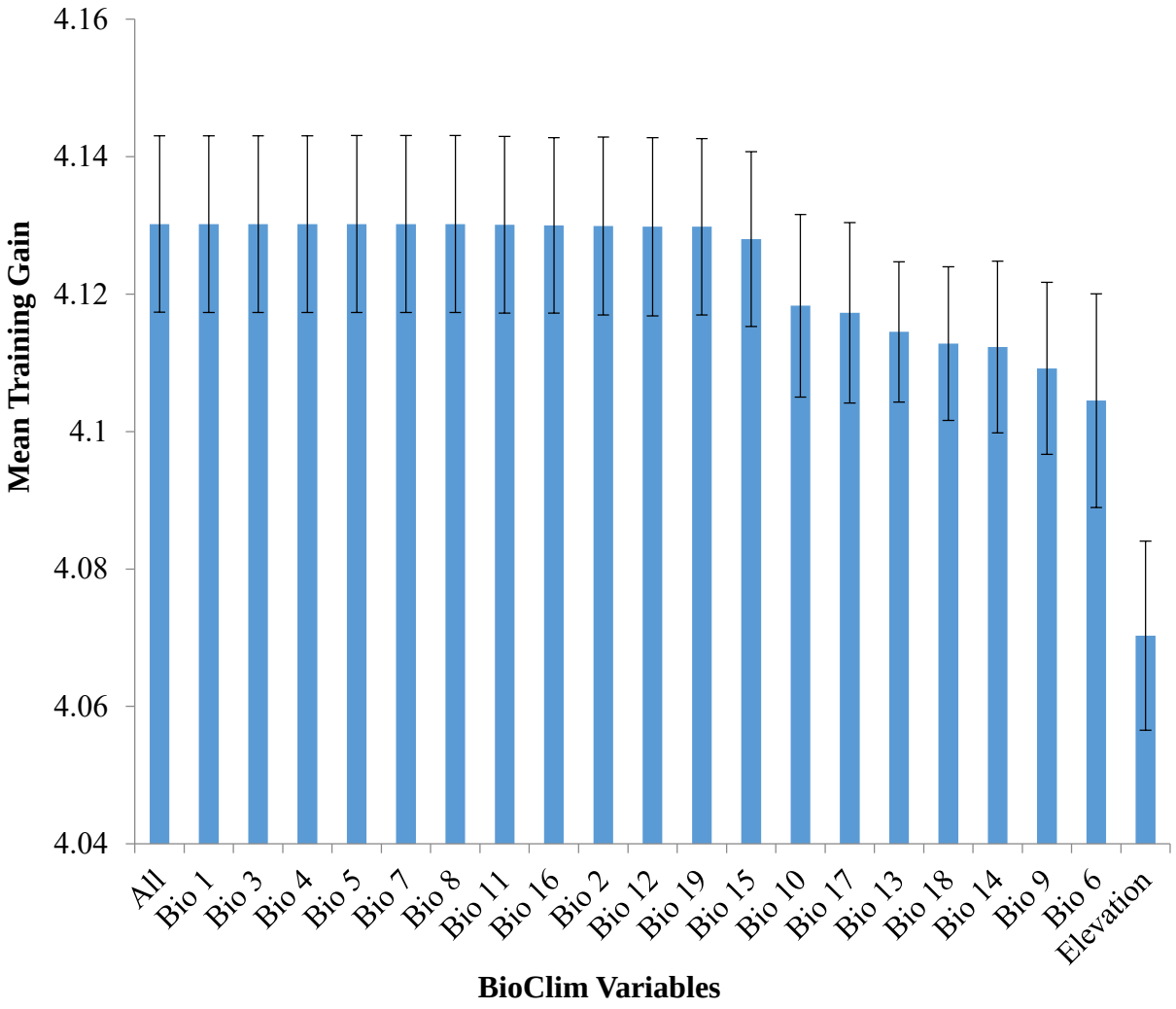

Figure 22: Jackknife test of variable importance with mean training gain (over 10 runs) for all variables and with each variable omitted. The variable that decreases the mean training gain the most when it is omitted and appears to have the most information that is not present in the other variables is elevation. BioClim variables correspond to Table 2.

Table 7: ENMTools model selection. Ecogeographical Variable (EGV) Combinations are all of the possible combinations of uncorrelated environmental variables from the nine variables selected as most important in model selection. Log Likelihood is the likelihood that observed hybrid presence points occur within the model. Sample Size is the number of hybrid locations. AIC (Akaike Information Criterion) Score is a measure of model quality that deals with the trade-off between goodness of fit and complexity of the model. AICc Score is AIC for models with a small sample size. Elevation and Bio 5 performed the best at modeling the *Archilochus* hybrid distribution (AICc = 390.814). Bio variables correspond to Table 2.

EGV Combinations	Log Likelihood	Sample Size	AIC Score	AICc Score
Elevation, Bio 5	-187.91	17	385.814	390.814
Elevation, Bio 10	-186.79	17	385.585	393.221
Elevation, Bio 9	-191.43	17	394.857	402.494
Elevation, Bio 11	-194.52	17	401.04	408.676
Elevation, Bio 6	-197.81	17	405.619	410.619
Elevation, Bio 1	-187.37	17	392.736	415.236
Bio 5	-209.42	17	422.84	423.64
Bio 9	-211.32	17	426.639	427.439
Bio 10	-212.25	17	428.507	429.307
Elevation, Bio 8	-205.25	17	422.501	430.137
Elevation, Bio 3	-204.92	17	423.834	435.034
Bio 1	-217.67	17	439.333	440.133
Bio 11	-219.36	17	442.72	443.52
Bio 6	-220.26	17	444.529	445.329
Bio 8	-220.28	17	446.551	448.265
Bio 3	-230.11	17	466.221	467.935
Elevation	-245.83	17	501.656	506.656

Internal statistical analysis done using MaxEnt suggested a good balance between sensitivity and specificity (AUC = 0.995 ± 0.004). The environmental variable BIO 5 (max temperature in warmest month) had the most useful and unique information for model generation. The MaxEnt habitat suitability map (Figure 23) revealed ~374,896 hectares of suitable habitat (suitability > 0.50) from central to western Oklahoma that tapers down to a narrower area in central Texas and curves around the south side of the Edwards Plateau ecoregion in Texas. Suitability for hybrids appears to be moderate in Arkansas, Missouri, and Kansas. The Mississippi River basin and Gulf Coast had very low habitat suitability (< 0.0001), which could be the result of having elevation as one of the utilized geographical variables.

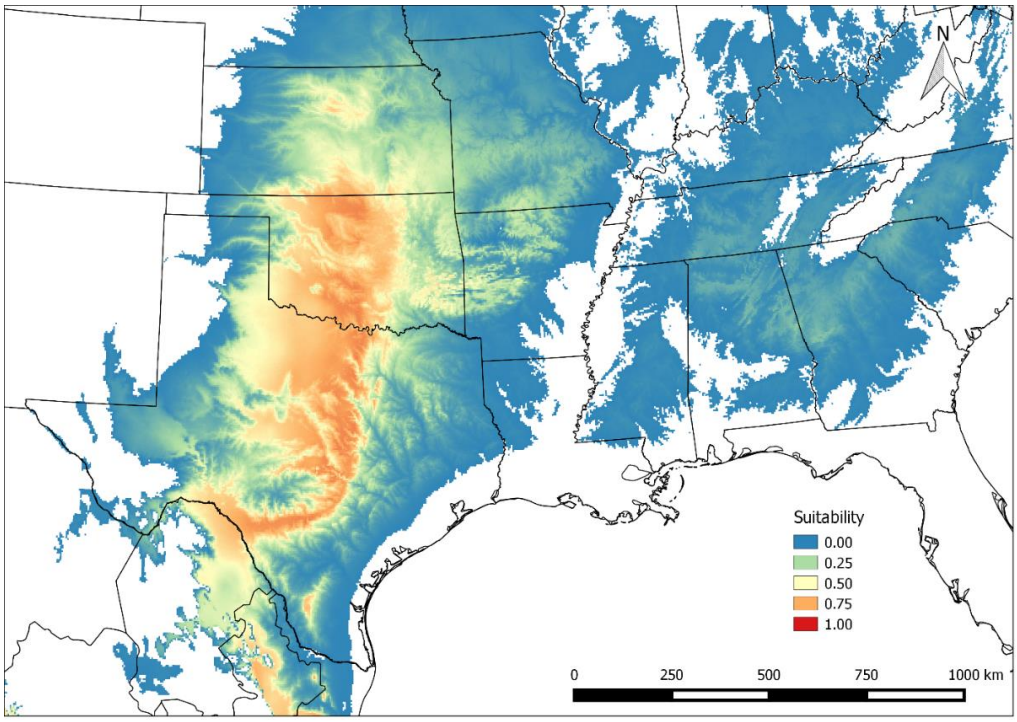

Figure 23: Point-wise mean of 10 habitat suitability maps calculated using MaxEnt v. 3.3.3k (Phillips et al. 2004, 2006). Red and orange indicate higher predicted habitat suitability for hybrids, yellow indicates conditions typical of those where hybrids were found, and blue indicates low predicted probability of suitable conditions for hybrids. Transparent areas are very low suitability (< 0.0001).

Internal measures of accuracy for the GARP model suggested a good balance between sensitivity and specificity as well (AUC = 0.986). Since GARP allows a moderate to high commission error while minimizing omission error, it is more likely to overestimate suitable habitat (Elith et al. 2006). GARP predicted ~956,441 hectares of suitable habitat (suitability > 0.50) for hybrids from western and central Oklahoma south to central Texas and west to New Mexico (Figure 24). This model had the greatest predicted amount of suitable habitat, which was consistent with the openModeller model selection algorithm (Anderson et al. 2003). The sharp delineation between very low suitability (transparent) and high suitability (red) is likely an artifact of the minimization of omission error.

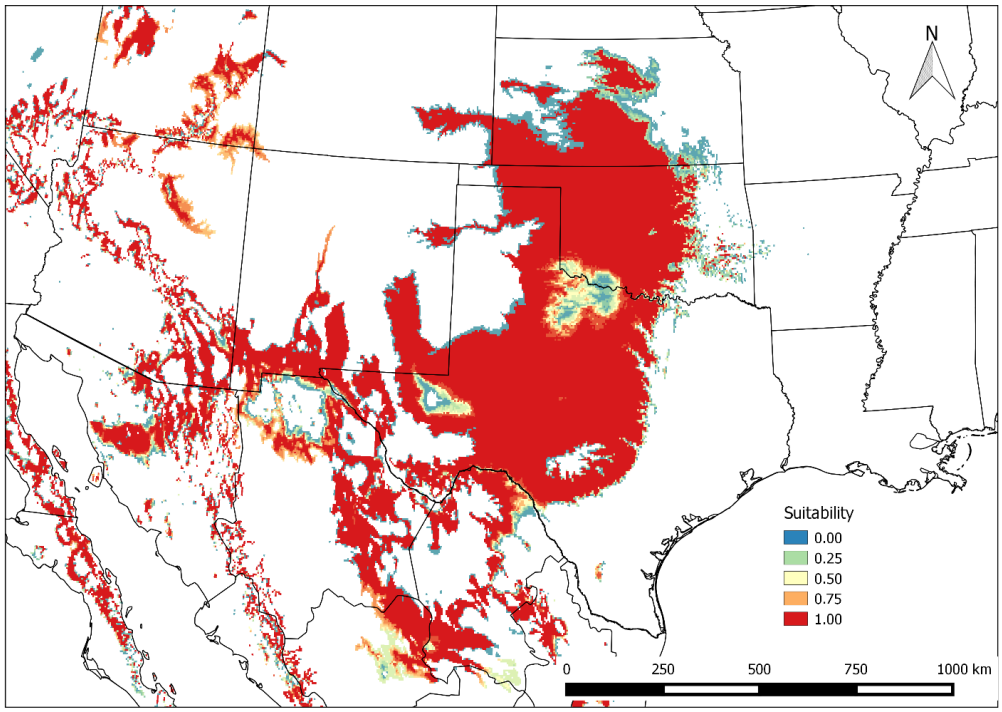

Figure 24: GARP with Best Subsets (new openModeller implementation) habitat suitability map as generated using openModeller Desktop v.1.1.0 (Muñoz et al. 2009). Blues and greens indicate lower probability of suitable habitat and yellows, oranges, and reds indicate higher habitat suitability for hybrids. Transparent areas are very low suitability (< 0.0001).

Model accuracy tests within Biomapper use a continuous Boyce index which ranges from 0 – 1, with 1 having the best fit to hybrid locations. The best model generated by ENFA had a continuous Boyce index of 0.408 ± 0.5, which is indicative of a poor fit. ENFA revealed that BIO 5 explained most of the variation in environmental factors between the presence points and the study area (92%). ENFA predicted ~364,600 hectares of suitable habitat (suitability > 0.50) from north central Oklahoma south to central Texas with a thin curve that wraps around the south side of the Edwards Plateau (Figure 25). ENFA also predicts areas of lower suitability that extend eastward to Florida, Georgia, and South Carolina. ENFA predicted the smallest area of suitable habitat for hybrids (with a suitability above 0.5).

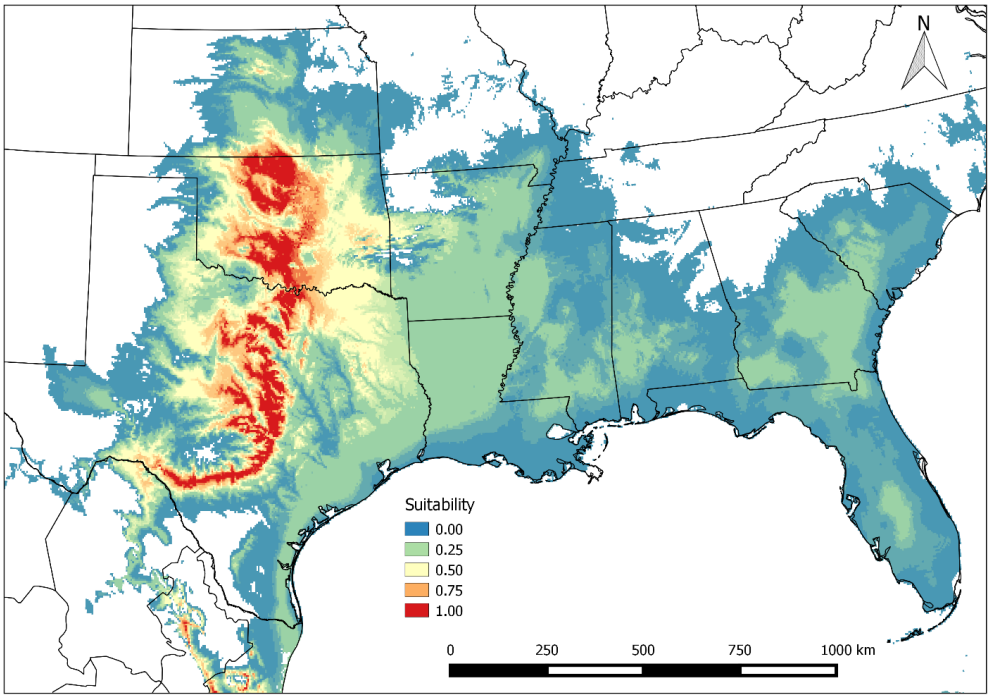

Figure 25: ENFA habitat suitability map as generated with the ecological geometric mean algorithm in Biomapper 4 (Hirzel et al. 2002b). White, blue, and green areas have a low suitability score, whereas yellow, orange, and red areas are more suitable for hybrids. Transparent areas have a very low probability of having suitable habitat for hybrids (< 0.0001).

Model accuracy statistics revealed that, at a threshold value of 0.5, GARP had the highest AUC (0.93), lowest omission rate (0.11), highest sensitivity (0.89), lowest specificity (0.98), and lowest proportion correct (0.97, Table 8). ENFA had the lowest AUC (0.74), highest omission rate (0.50), lowest sensitivity (0.50), highest specificity (0.99), and highest proportion correct (0.99). MaxEnt had the highest Kappa, but medial values in every other measure of accuracy. Comparisons of model accuracy statistics among the three ecological niche models should be interpreted with caution because the models did not take into account where the current breeding ranges overlap, only where the genetic introgression occurred.

Table 8: Tests for ecological niche model accuracy as estimated using the SDMTools Package v. 1.1-221 for R (VanDerWal et al. 2014). Area Under the Curve (AUC; threshold independent accuracy measure), Omission Rate (proportion of true occurrences misidentified), Sensitivity (proportion of true positive predictions), Specificity (proportion of false positive predictions), Proportion Correct (proportion of predicted presences that were correct), and Kappa (threshold dependent accuracy measure) were calculated at a threshold value of 0.5. All measures of accuracy range from 0 to 1. A good model has a high AUC, low omission rate, high sensitivity and specificity values (so one is not sacrificed to achieve the other), high proportion of correct predictions, and a high Kappa.

	Threshold	AUC	Omission Rate	Sensitivity	Specificity	Proportion Correct	Kappa
ENFA	0.5	0.74452	0.5000	0.5000	0.9890	0.9890	0.00083
GARP	0.5	0.93006	0.1111	0.8889	0.9847	0.9712	0.00056
MaxEnt	0.5	0.91102	0.1667	0.8333	0.9887	0.9887	0.00137

Chapter IV

DISCUSSION

This study had four goals. The first goal was to identify hybrids genetically. The second goal was to determine the extent of the *Archilochus* hybrid zone with ecological niche modeling. The third goal was to compare the accuracy of three ecological niche model habitat suitability maps. The final goal was to use the most accurate habitat suitability map to determine which hybrid zone theory best describes the *Archilochus* hybrid zone.

Genetic Identification of Hybrids

Dispersal of Mitochondrial Types

Sequencing mtDNA allowed for identification of maternal species in some individuals. Tavares and Baker (2008) described parental Black-chinned Hummingbird barcode types (TCTGG) in NM and AZ, and I was able to identify them in the Wichita Mountains Wildlife Refuge (Comanche County) and Chickasha (Grady County) in southwest OK. A new Black-chinned Hummingbird type (CCTAA) was identified in the established Black-chinned Hummingbird breeding range, as well as new dispersal areas in southwest and central OK (see Butler et al. 2007). Black-chinned Hummingbird type (CCTAA) sites were located from north of Guthrie (Logan County), OK southwest to Christoval (Tom Green County), TX and then southeast to Leakey (Real County), TX and Austin (Travis County), TX and finally north to Ft. Worth (Tarrant County), TX and Arcadia (Oklahoma County), OK.

Tavares and Baker (2008) identified the Ruby-throated Hummingbird barcode type (CACAA) in southeast Ontario (Canada) and Nova Scotia (Canada). I was able to identify this type in Alabama, Arkansas, Oklahoma, and Texas as far west as Christoval. The proportion of

the Ruby-throated Hummingbird type, in the examined populations, appears to decline west of the established Ruby-throated Hummingbird breeding zone.

The new Black-chinned Hummingbird type (CCTAA) is one of four intermediate types between Black-chinned Hummingbirds and Ruby-throated Hummingbirds. Two of the other intermediates (Haplotype 19 – CATAA, Haplotype 28 – TCTAA) occur only in Chickasha, OK. The remaining intermediate type was found close to Ponca (Newton County), AR (Haplotype 69 – CCCAA). Based on the locations of these intermediates, it appears that there is a mtDNA cline across the south-central US. This is consistent with the assumption made by Barton and Hewitt (1985) in which a genetic gradient exists across the hybrid zone landscape. Further analysis of *Archilochus* populations in the US would reveal whether these initial findings are due to isolated mutations or an overall geographic trend.

Phylogenetic Trees

The greatest variability in haplotypes (entire sequence) was observed in the Great Plains. Only three haplotypes were observed in Black-chinned Hummingbird individuals from NM and AZ, and only four haplotypes were observed in Ruby-throated Hummingbird individuals from Canada. Small sample sizes in the Tavares and Baker (2008) study could have affected the diversity of haplotypes observed, but the lack of genetic variation could also indicate genetic bottlenecks on each side of the hybrid zone. Attempts to identify maternal species by the construction of phylogenetic trees were unsuccessful based on low bootstrap values at most nodes (Figures 17 and 18). In phylogenies generated by Tavarez and Baker (2008) and McGuire et al. (2007, 2014), the Black-chinned Hummingbird and Ruby-throated Hummingbird phenotypes occurred on separate branches of the trees. As expected in a hybrid zone, phenotypes in this study did not group with each other consistently.

Low bootstrap values reflect the low confidence in relationships at most nodes. Short branch lengths suggest low overall sequence divergence, which is expected in mtDNA of closely related species. Intraspecific phylogenies are difficult to infer by traditional tree-building techniques because they violate assumptions that are made by the algorithms (Crandall et al. 1994). Individuals with mitochondrial barcoding types that were inconsistent with Tavares and Baker (2008) caused more species-level clusters than expected. Individuals that had ambiguous base calls contributed to the number of unique haplotypes.

Clusters that were observed seem to suggest parental populations that are not completely divergent. This also is supported by the presence of intermediate haplotypes from west TX and OK to central TX and OK, and by a gradual tapering off of Ruby-throated Hummingbird haplotypes west of the contact zone (well into the Black-chinned Hummingbird breeding range).

Population Structure Analysis

The three microsatellite loci tested were not sufficient to group individuals in a way that assisted in hybrid identification. However, addition of mtDNA SNP loci to the microsatellite data provided enough variability to identify two clusters of individuals that were consistent with Black-chinned Hummingbird and Ruby-throated Hummingbird breeding ranges. Support for these clusters was high for most individuals; however, there were a couple of cases where the confidence intervals did not overlap but were very broad. Caution should be used when interpreting these cluster assignments because of the size of each confidence interval.

Clusters were not consistent with Tavares and Baker (2008) barcode types. The Ruby-throated Hummingbird barcode type (CACAA) was identified as one cluster, but the most common intermediate type (CCTAA) was identified as the other cluster. The two individuals with Black-chinned Hummingbird barcode types (TCTGG) were identified as a mixture of the

two clusters. This makes sense because there were so few Black-chinned Hummingbird barcode types and many more intermediate types in the established Black-chinned Hummingbird breeding range. The CCCAA barcode type clustered with Ruby-throated Hummingbird, but the TCTAA and CATAA barcode types were identified as a mixture of the two clusters. There were several cases where the Structure identification was labeled as uncertain due to mixed cluster assignment, and there were a few cases where uncertainty was based on overlapping confidence intervals.

Overall, a broad area of introgression was found to exist from west Texas to central Texas and central Oklahoma. This was farther west than expected because apparent hybrid males have only been documented as far west as Chickasha, OK.

Implications for Hybrid Zone History

Hybridization can occur during parapatric speciation or following allopatric speciation, the difference in the two is timing of hybridization. Allopatric hybridization occurs after geographically separate, genetically diverged populations come back into contact and interbreed. Parapatric hybridization occurs during the speciation process, without geographic separation, due to differences in environment. Previously reported observations in the literature of two mitochondrial types (distinct species, Tavares and Baker 2008), individuals with pink gorgets (hybridization, Judd et al. 2011), and recent northward expansion of Black-chinned Hummingbirds (Butler et al. 2007) suggested allopatric hybridization, where Black-chinned Hummingbirds and Ruby-throated Hummingbirds were hybridizing in a relatively recent secondary contact zone. However, mitochondrial DNA in Oklahoma and Texas did not strongly support fixed differences in the Black-chinned Hummingbird and Ruby-throated Hummingbird haplotypes, as would be expected if separation had occurred. Tavares and Baker (2008)

identified five COI SNPs that were fixed for each of the *Archilochus* species they examined (identifying barcode types). These same loci in the Oklahoma and Texas samples had intermediates between the two species, suggesting a possible mtDNA cline and thus, parapatric speciation. Hybridization is common with parapatric speciation, but the point at which intraspecific variation stops and hybridization begins is difficult to determine.

Implications for Archilochus Radiation in the Bee Hummingbird Clade

McGuire et al. (2014) published a time-calibrated phylogenetic tree of the Trochilidae family and it suggested that the *Archilochus* genus diverged between 2.5 and 2.0 million years ago in North America. This places diversification during the Pleistocene when multiple glaciation events could have affected the genetic diversity of *Archilochus*. If climate shifts during glaciation caused barriers to gene flow in the ancestral *Archilochus* populations, genetic drift would inevitably cause random alleles to drop out of each population. According to speciation theory, once genetic differences reach the point where interbreeding cannot occur, speciation is complete (Endler 1977). Speciation caused by glaciation typically produces sharp bursts of divergence that are tied to the appearance of glaciers.

The bee hummingbird clade (which includes *Archilochus*) had a sharp burst in diversification rate (and then sharp decline) between 5 mya (million years ago) and present (McGuire et al. 2014) that coincided with the Cascade/Sierra chain uplift. The new mountains cast a rain shadow that caused xerification of the Columbia Plateau (Graham 1999). Pleistocene glaciation (in 100,000 year cycles) between 2.5 mya and 11,000 years ago had enormous impacts on the geographic distributions of organisms in the region (Brunsfeld et al. 2001). Between 12 and 5 mya, the bee hummingbird clade was relatively stable and had no apparent species diversification, while the mountain gem clade was rapidly diversifying (McGuire et al. 2014). It

is possible that the common ancestor to bee hummingbirds was widespread across the US during this stable period and the Cascade/Sierra uplift jump started the bee clade divergence. Glacial events during the Pleistocene further changed the environment and *Archilochus* could have diverged as a result of newly available niches.

The location in North America where *Archilochus* originally diverged could reveal more about how the Ruby-throated Hummingbird and Black-chinned Hummingbird phenotypes developed. If one type was the result of a genetic bottleneck, then a hub of greater genetic diversity should be present in one species breeding range, with diversity decreasing as you travel toward the other species (Nei et al. 1975). If both species diverged from a common ancestor that is no longer genetically represented, then the highest genetic diversity for both species should center on, or be close to, the ancestral divergence site. In this case, diversity in both species appears to decrease as you move away from the hybrid zone. The area with the greatest genetic diversity appears to be at the contact zone between the two species. Haplotypes were present in the contact zone that were not observed in either species outside of the contact zone (including in the zone of introgression). Since this area of greatest diversity occurs at the species boundary, it is possible that the Ruby-throated Hummingbird and the Black-chinned Hummingbird diverged around the same time in response to changes in the environment that created available niches. This cannot be confirmed in this study because more sites in the eastern and western US need to be sampled.

Ecological Niche Modeling

Ecogeographical Variables

Elevation and max temperature of warmest month (Bio 5) were most successful at creating an accurate niche model for hybrids. However, elevation and mean temperature of warmest quarter (Bio 10) produced a model with an AIC score that was very close to the elevation and Bio 5 model AIC. It is evident that elevation was important in modeling the distribution of hybrids, but it is unclear whether mean temperature in summer or the maximum temperature in the warmest month was most important.

Elevation, which was one of the important factors in the hybrid niche models, transitions along the phenotype contact zone in Texas. Elevation differences in the Andes have been associated with speciation in hummingbirds based on efficiency of hemoglobin at different elevations (Projecto-Garcia et al. 2013). It is possible that the phenotype boundary has been maintained by metabolic advantages needed to thrive in the higher elevations; however, the difference in elevation at the *Archilochus* phenotype boundary is not as great as in the Andes, so a metabolic advantage is unlikely.

Model Comparison

All three models predicted that the hybrid zone is characterized by a broad north/south band of suitable habitat from central Oklahoma to south-central Texas. This was consistent with the extent of genetic introgression in western and central Oklahoma and Texas.

Area under the curve (AUC) should be high (close to 1) in a good model. GARP had the highest AUC, which suggests that it had the best balance between sensitivity and specificity. However, AUC vales should be interpreted cautiously, because they can be misleading (Peterson et al. 2007).

Sensitivity (the proportion of correctly classified presences) and specificity (proportion of correctly classified absences) should both be high in a good model. All three models had high specificity, wheras only GARP had a relatively high sensitivity. The high sensitivity achieved by GARP was likely the result of over-prediction of suitable habitat. The habitat suitability map had a large area of high probability of hybrids. This increased the chance that a presence would be classified correctly because there was no variability in habitat suitability. The high specificity value for all three models is most likely misleading. True absence data points were not available for this study, and the model assumed that any pixel that did not have a presence point was an absence. The likelihood of encountering an absence was much higher than encountering a presence point, so very high specificity rates were possible. Given true absence data, specificity would probably be more variable for the different models.

The omission rate should be low for a good model performance. It indicates the percentage of test localities that fall into pixels not predicted as suitable for the species (Phillips et al. 2006). The low omission rate that GARP achieved was likely due to an inherent characteristic of the model, which is to avoid omission errors.

The proportion correct (the percentage of sites where the model correctly predicted either presence or absence) should be high for a good model. All three models had a high proportion of sites that had correct predictions. This was likely related to the high proportion of sites that were labeled as absences. True absence locations could change these results.

Comparison of the accuracy measurements of the three ecological niche models revealed that the predictive power of MaxEnt and GARP was approximately the same, and that they were both greater than ENFA. This corresponds to the study done by Peterson et al. (2007) in which they compared the predictive power of MaxEnt and GARP. However, this similarity in

predictive power is achieved in different ways. MaxEnt avoids commission errors (false positives) which could include areas in the habitat suitability map that are not suitable for the species (Ward 2007). GARP avoids omission errors (false negatives) which could leave out areas in the habitat suitability map where the species exists (Anderson et al. 2003).

Applicability to Hybrid Zones

Niche modeling the *Archilochus* hybrid zone was moderately successful, but the predicted habitat suitability for hybrids was larger than the parental breeding range overlap. By definition, hybrids occur where parental species come into contact and interbreed, so habitat suitability was likely overestimated. Including true absence sites in the model would likely increase the validity of each model's accuracy measures. It is difficult to attempt to isolate factors that affect hybrid distributions because hybrids are often dependent upon the location and overlap of their parental populations. However, it can be helpful in determining whether a hybrid zone is dependent upon environmental factors.

Hybrid Zone Theory

The recent expansion of Black-chinned Hummingbirds into Ruby-throated Hummingbirds breeding range in Oklahoma (Butler et al. 2007) suggests that the northeastern part of the hybrid zone is shifting. This suggests that the transient/ephemeral hybrid zone theory applies to this section of the *Archilochus* hybrid zone. However, the data in this study only represents a snapshot in time. Further analysis of the hybrid zone over time could reveal whether one species is replacing the other in the recent area of Black-chinned Hummingbird expansion.

The bounded hybrid superiority model seems to fit the southwest Oklahoma and the north central Texas portion of the contact zone as well because hybrids outnumber parental types in a

few areas. This could indicate a competitive advantage that the hybrids have over the pure parental types. There could be slight variations in plumage color or wing shape of hybrids based on preliminary analysis of photos taken; however, the subtle differences were not quantified during specimen handling.

Ecological niche modeling suggested that the hybrid zone is affected by environmental conditions. This contradicts the dynamic-equilibrium/tension zone model assumption that the hybrid zone is not affected by environmental conditions (Barton and Hewitt 1989). Genetic analysis revealed that the zone of introgression into the Black-chinned Hummingbird phenotype breeding zone is broad. This also contradicts the dynamic-equilibrium/tension zone model in which hybrids are found only in a narrow contact zone. Despite the evidence against the tension zone model, if the population dynamics east of the contact zone are observed separate from the western zone of introgression, evidence of strong selection against certain hybrids emerges. Only two individuals were observed in the contact zone that were field identified as Ruby-throated Hummingbird but genetically identified as Black-chinned Hummingbird. East of the contact zone, there was no evidence of Black-chinned Hummingbird introgression into Ruby-throated Hummingbird populations. It appears that different hybrid zone theories could fit different aspects of the *Archilochus* hybrid zone; however, more sample locations and genotype data is needed to assess this point.

The Ruby-throated Hummingbird mtDNA introgression into the Black-chinned Hummingbird phenotype area suggests that hybridization is occurring where the different phenotypes come into contact, but then the Ruby-throated Hummingbird phenotype is selected against west of the contact zone. Ruby-throated Hummingbird mtDNA continues to be passed from mother to offspring despite the dilution of Ruby-throated Hummingbird nuclear genes;

thus, a population can consist of Black-chinned Hummingbird phenotypes, but still contain Ruby-throated Hummingbird mtDNA. The lack of Black-chinned Hummingbird mtDNA and phenotypes east of the contact zone suggests that selection against Black-chinned Hummingbirds is strong. The selection that occurs seems to contradict the hybrid-equilibrium model that assumes that there is no selection against hybrids or parent organisms (Van Den Bussche et al. 1993).

Conclusions

The identification of hybrids using mitochondrial DNA revealed that there is more genetic introgression than previously thought. Because hybrid identification depended upon each parental species having a unique mitochondrial type, and more than two types were observed, further analysis of parental mtDNA should be done. All three ecological niche models used in this study predict a similar hybrid range that was broad throughout Oklahoma and Texas. Variability in predicted area of high habitat suitability allowed for direct comparison of accuracy. When only accuracy statistics are considered, GARP appeared to perform best at hybrid zone modelling. However, previous studies have noted that GARP tends to over-estimate suitable habitat. Optimization of each model would improve individual accuracy before comparison. The *Archilochus* hybrid zone appears to be stable, and the greatest amount of genetic diversity is at the center of the hybrid zone. This suggests that the bounded hybrid superiority model best describes the *Archilochus* hybrid zone; however, recent expansion of Black-chinned Hummingbirds into Oklahoma suggests transience. It is possible that the older contact zone in central Texas is best described by bounded hybrid superiority and the new expansion is more characteristic of an ephemeral hybrid zone.

Future Research

Sampling of Black-chinned Hummingbirds further west of Christoval, TX is necessary to determine if a unique mitochondrial type exists for Black-chinned Hummingbirds. If unique mitochondrial types do not exist for Black-chinned Hummingbird, then additional mitochondrial genes should also be sequenced to assist in maternal lineage analysis. Several microsatellite loci also need to be developed for the analysis of population structure within *Archilochus.* Optimization of each ecological niche model before comparison should be done in order to utilize their individual strengths, and genetic sampling of *Archilochus* over time would address whether the hybrid zone is stable or ephemeral.

APPENDICES:

Appendix A: Hummingbirds captured. Sample # corresponds to numbers on Structure output (Figure 19); Band number is the Bird Banding Lab leg band number; Field IDs are alpha codes where RTHU = Ruby-throated Hummingbird, BCHU = Black-chinned Hummingbird, and UNHY = Unknown Hybrid; Latitude and Longitude indicate location of capture.

Sample #	Band Number	Date Collected	Field ID	Latitude	Longitude
1	3100H-35801	7/17/2011	RTHU	33.724453	-86.874531
2	3100H-99288	7/17/2011	RTHU	33.724453	-86.874531
3	3100H-99293	7/17/2011	RTHU	33.724453	-86.874531
4	3100H-99512	7/17/2011	RTHU	33.724453	-86.874531
5	4100L-68624	7/17/2011	RTHU	33.724453	-86.874531
6	4100L-68626	7/17/2011	RTHU	33.724453	-86.874531
7	4100L-68631	7/17/2011	RTHU	33.724453	-86.874531
8	4100L-68635	7/17/2011	RTHU	33.724453	-86.874531
9	9000E-57680	7/17/2011	RTHU	33.724453	-86.874531
10	9000E-57685	7/17/2011	RTHU	33.724453	-86.874531
11	E44516	6/19/2010	BCHU	35.058692	-97.97195
12	H25244	6/7/2010	RTHU	35.732031	-97.289144
13	H25245	6/7/2010	RTHU	35.732031	-97.289144
14	H25246	6/7/2010	RTHU	35.732031	-97.289144
15	H25247	6/7/2010	RTHU	35.732031	-97.289144
16	H25248	6/7/2010	RTHU	35.732031	-97.289144
17	H25249	7/8/2010	BCHU	35.058692	-97.97195
18	H25250	7/9/2010	BCHU	34.731764	-98.712008
19	H25251	7/9/2010	BCHU	34.731764	-98.712008
20	H25252	7/9/2010	BCHU	34.731764	-98.712008
21	H25256	7/15/2010	BCHU	35.058692	-97.97195
22	H25257	7/15/2010	BCHU	35.058692	-97.97195
23	H25259	7/15/2010	BCHU	35.058692	-97.97195
24	H25260	7/15/2010	BCHU	35.058692	-97.97195

Appendix A: Continued

Sample #	Band Number	Date Collected	Field ID	Latitude	Longitude
25	H25330	6/13/2010	BCHU	34.631583	-98.377864
26	H25331	6/13/2010	BCHU	34.631583	-98.377864
27	H25334	6/13/2010	BCHU	34.631583	-98.377864
28	H25335	6/19/2010	RTHU	35.058692	-97.97195
29	H25336	6/19/2010	BCHU	35.058692	-97.97195
30	H25337	6/20/2010	BCHU	34.779222	-98.606111
31	H25338	6/20/2010	BCHU	34.779222	-98.606111
32	H25339	6/20/2010	BCHU	34.779222	-98.606111
33	H25340	6/20/2010	BCHU	34.779222	-98.606111
34	H25341	6/20/2010	BCHU	34.779222	-98.606111
35	H25342	6/20/2010	BCHU	34.779222	-98.606111
36	H25343	6/20/2010	BCHU	34.779222	-98.606111
37	H25345	6/24/2010	RTHU	35.606964	-97.608853
38	H25346	6/26/2010	BCHU	34.63195	-98.722461
39	H25347	6/26/2010	BCHU	34.63195	-98.722461
40	H25348	6/26/2010	BCHU	34.63195	-98.722461
41	H25349	6/26/2010	BCHU	34.63195	-98.722461
42	H25350	6/26/2010	BCHU	34.63195	-98.722461
43	H25351	6/29/2010	BCHU	30.442361	-97.815097
44	H25352	6/29/2010	BCHU	30.442361	-97.815097
45	H25353	6/29/2010	BCHU	30.442361	-97.815097
46	H25354	6/29/2010	BCHU	30.442361	-97.815097
47	H25355	6/29/2010	BCHU	30.442361	-97.815097
48	H25356	6/30/2010	BCHU	30.488008	-98.055411
49	H25357	6/30/2010	BCHU	30.488008	-98.055411
50	H25358	6/30/2010	BCHU	30.488008	-98.055411
51	H25359	6/30/2010	BCHU	30.623267	-98.104536
52	H25360	6/30/2010	BCHU	30.623267	-98.104536
53	H25361	7/1/2010	BCHU	30.349906	-98.058606

Appendix A: Continued

Sample #	Band Number	Date Collected	Field ID	Latitude	Longitude
54	H25362	7/1/2010	BCHU	30.349906	-98.058606
55	H25363	7/1/2010	BCHU	30.349906	-98.058606
56	H25364	7/1/2010	BCHU	30.349906	-98.058606
57	H25365	7/1/2010	BCHU	30.349906	-98.058606
58	H25366	7/1/2010	BCHU	30.349906	-98.058606
59	H25368	7/1/2010	BCHU	30.349906	-98.058606
60	H25369	7/1/2010	BCHU	30.349906	-98.058606
61	H25370	7/1/2010	BCHU	30.349906	-98.058606
62	H25371	7/1/2010	BCHU	30.349906	-98.058606
63	H25372	7/1/2010	BCHU	30.349906	-98.058606
64	H25373	7/1/2010	BCHU	30.349906	-98.058606
65	H25374	7/1/2010	BCHU	30.349906	-98.058606
66	H25375	7/1/2010	BCHU	30.349906	-98.058606
67	H25376	7/1/2010	BCHU	30.180678	-97.845753
68	H25379	7/2/2010	BCHU	30.303119	-98.087956
69	H25380	7/2/2010	BCHU	30.303119	-98.087956
70	H25381	7/2/2010	BCHU	30.303119	-98.087956
71	H25382	7/2/2010	BCHU	30.303119	-98.087956
72	H25383	7/2/2010	BCHU	30.303119	-98.087956
73	H25384	7/2/2010	BCHU	30.303119	-98.087956
74	H25385	7/2/2010	BCHU	30.303119	-98.087956
75	H25386	7/2/2010	BCHU	30.303119	-98.087956
76	H25387	7/2/2010	BCHU	30.303119	-98.087956
77	H25388	7/2/2010	BCHU	30.303119	-98.087956
78	H25389	7/2/2010	BCHU	32.1306	-97.854822
79	H25390	7/2/2010	BCHU	32.1306	-97.854822
80	H25391	7/2/2010	BCHU	32.1306	-97.854822
81	H25392	7/2/2010	BCHU	32.1306	-97.854822
82	H25393	7/2/2010	BCHU	32.1306	-97.854822

Appendix A: Continued

Sample #	Band Number	Date Collected	Field ID	Latitude	Longitude
83	H25394	7/2/2010	BCHU	32.1306	-97.854822
84	H25395	7/2/2010	BCHU	32.1306	-97.854822
85	H25396	6/7/2010	RTHU	35.732031	-97.289144
86	H25397	6/7/2010	RTHU	35.732031	-97.289144
87	H25398	6/7/2010	RTHU	35.732031	-97.289144
88	H25399	6/7/2010	RTHU	35.732031	-97.289144
89	H25400	6/7/2010	RTHU	35.732031	-97.289144
90	L34317	6/26/2010	BCHU	29.767169	-99.723506
91	L34341	6/26/2010	BCHU	29.897111	-98.864164
92	L34342	6/26/2010	BCHU	29.897111	-98.864164
93	L34344	6/26/2010	BCHU	29.897111	-98.864164
94	L34350	6/26/2010	BCHU	29.897111	-98.864164
95	L34351	6/26/2010	BCHU	29.897111	-98.864164
96	L34403	6/26/2010	BCHU	29.767169	-99.723506
97	L34404	6/26/2010	BCHU	29.767169	-99.723506
98	L34405	6/26/2010	BCHU	29.767169	-99.723506
99	L34410	6/26/2010	BCHU	29.767169	-99.723506
100	L34414	6/26/2010	BCHU	29.767169	-99.723506
101	L34415	6/26/2010	BCHU	29.767169	-99.723506
102	L34910	6/26/2010	BCHU	29.767169	-99.723506
103	L34922	6/26/2010	BCHU	29.767169	-99.723506
104	L35651	6/26/2010	BCHU	29.897111	-98.864164
105	L35652	6/26/2010	BCHU	29.897111	-98.864164
106	L35653	6/26/2010	BCHU	29.897111	-98.864164
107	L35654	6/26/2010	BCHU	29.897111	-98.864164
108	L4100-68539	6/18/2011	BCHU	31.164074	-100.496986
109	L4100-68541	6/18/2011	BCHU	31.164074	-100.496986
110	L4100-68542	6/18/2011	BCHU	31.164074	-100.496986
111	L4100-68543	6/18/2011	BCHU	31.164074	-100.496986

Appendix A: Continued

Sample #	Band Number	Date Collected	Field ID	Latitude	Longitude
112	L4100-68544	6/18/2011	BCHU	31.164074	-100.496986
113	L4100-68545	6/18/2011	BCHU	31.164074	-100.496986
114	L4100-68546	6/18/2011	BCHU	31.164074	-100.496986
115	L4100-68549	6/18/2011	BCHU	31.164074	-100.496986
116	L4100-68550	6/18/2011	BCHU	31.164074	-100.496986
117	L4100-68551	6/18/2011	BCHU	31.164074	-100.496986
118	L60380	6/1/2011	BCHU	33.385294	-97.489114
119	L60382	6/1/2011	BCHU	33.385294	-97.489114
120	L60383	6/1/2011	BCHU	33.385294	-97.489114
121	L60384	6/7/2011	UNHY	33.385294	-97.489114
122	L60386	6/7/2011	BCHU	33.344444	-97.465278
123	L60387	6/7/2011	BCHU	33.344444	-97.465278
124	L60388	6/7/2011	BCHU	33.344444	-97.465278
125	L60389	6/7/2011	BCHU	33.344444	-97.465278
126	L60391	6/7/2011	BCHU	33.344444	-97.465278
127	L60392	6/7/2011	BCHU	33.344444	-97.465278
128	L60393	6/7/2011	BCHU	33.344444	-97.465278
129	L60394	6/7/2011	BCHU	33.344444	-97.465278
130	L60395	6/7/2011	BCHU	33.344444	-97.465278
131	L60396	6/7/2011	BCHU	33.344444	-97.465278
132	L60397	6/7/2011	BCHU	33.328333	-97.667222
133	L60398	6/7/2011	BCHU	33.328333	-97.667222
134	L60399	6/7/2011	BCHU	33.328333	-97.667222
135	L60401	7/16/2011	BCHU	33.756106	-97.310997
136	L60403	7/16/2011	BCHU	33.756106	-97.310997
137	L60404	7/16/2011	BCHU	33.756106	-97.310997
138	L60405	7/16/2011	BCHU	33.756106	-97.310997
139	L60406	7/16/2011	BCHU	33.756106	-97.310997
140	L60407	7/16/2011	BCHU	33.756106	-97.310997

Appendix A: Continued

Sample #	Band Number	Date Collected	Field ID	Latitude	Longitude
141	L60408	7/16/2011	BCHU	33.756106	-97.310997
142	L60409	7/16/2011	BCHU	33.756106	-97.310997
143	L60410	7/16/2011	BCHU	33.756106	-97.310997
144	L60411	7/17/2011	BCHU	33.144786	-97.144192
145	L60412	7/17/2011	BCHU	33.144786	-97.144192
146	L60451	6/11/2011	RTHU	36.275814	-95.583217
147	L60452	6/11/2011	RTHU	36.275814	-95.583217
148	L60454	6/11/2011	RTHU	36.275814	-95.583217
149	L60456	6/11/2011	RTHU	36.275814	-95.583217
150	L60457	6/11/2011	RTHU	36.275814	-95.583217
151	L60458	6/11/2011	RTHU	36.275814	-95.583217
152	L60459	6/11/2011	RTHU	36.275814	-95.583217
153	L60460	6/11/2011	RTHU	36.275814	-95.583217
154	L60461	6/11/2011	RTHU	36.275814	-95.583217
155	L60462	6/12/2011	RTHU	35.997194	-93.393722
156	L60463	6/12/2011	RTHU	35.997194	-93.393722
157	L60464	6/12/2011	RTHU	35.997194	-93.393722
158	L60465	6/12/2011	RTHU	35.997194	-93.393722
159	L60466	6/12/2011	RTHU	35.997194	-93.393722
160	L60467	6/12/2011	RTHU	35.997194	-93.393722
161	L60468	6/12/2011	RTHU	35.997194	-93.393722
162	L60469	6/12/2011	RTHU	35.997194	-93.393722
163	L60470	6/12/2011	RTHU	35.997194	-93.393722
164	L60471	6/12/2011	RTHU	35.997194	-93.393722
165	L60472	6/19/2011	BCHU	33.36405	-97.804269
166	L60473	6/19/2011	BCHU	33.36405	-97.804269
167	L60474	6/19/2011	BCHU	33.36405	-97.804269
168	L60475	6/19/2011	BCHU	33.36405	-97.804269
169	L60476	6/19/2011	BCHU	33.36405	-97.804269

Appendix A: Continued

Sample #	Band Number	Date Collected	Field ID	Latitude	Longitude
170	L60477	6/24/2011	RTHU	35.694381	-97.542708
171	L60478	6/25/2011	RTHU	35.479067	-97.217103
172	L60479	6/25/2011	RTHU	35.479067	-97.217103
173	L60481	7/2/2011	RTHU	36.46395	-95.996175
174	L60482	7/2/2011	RTHU	36.46395	-95.996175
175	L60483	7/2/2011	RTHU	36.46395	-95.996175
176	L60484	7/2/2011	RTHU	36.46395	-95.996175
177	L60485	7/2/2011	RTHU	36.46395	-95.996175
178	L60486	7/2/2011	RTHU	36.46395	-95.996175
179	L60487	7/2/2011	RTHU	36.46395	-95.996175
180	L60488	7/2/2011	RTHU	36.46395	-95.996175
181	L60489	7/7/2011	RTHU	36.46395	-95.996175
182	L60490	7/3/2011	RTHU	35.997178	-97.223164
183	L60491	7/3/2011	RTHU	35.997178	-97.223164
184	L60492	7/14/2011	RTHU	35.997178	-97.223164
185	L60493	7/14/2011	RTHU	35.997178	-97.223164
186	L60494	7/14/2011	RTHU	35.997178	-97.223164
187	L60495	7/14/2011	RTHU	35.997178	-97.223164

Appendix B: Microsatellite and mtDNA SNP ancestry of individuals for two clusters (best run) as determined using Structure v. 2.3.4 (Pritchard et al. 2000). Band represents the Bird Banding Lab band number; % Mis. is the percentage of data missing for the individual; Inferred Clusters represents the proportion of an individual's analyzed genes that belong to each cluster; 95% Probability Intervals are the Bayesian equivalent of a confidence interval. Field ID is the species identification based on plumage (R = Ruby-throated, B = Black-chinned, U = unknown hybrid). Structure ID is species identification based on Structure clusters (R = Ruby-throated, B = Black-chinned, Uncertain = lacks statistical support). Haplotype is a number assigned to individuals with identical mtDNA sequences and corresponds to haplotypes in Table 5. Samples in bold have conflicting field identifications and structure identifications.

Band	% Mis.	Inferred Clusters		95% Probability Intervals		Field ID	Structure ID	Haplotype
3100H-35801	0	1.00	0.00	0.958,1.000	0.000,0.042	R	R	9
3100H-99288	0	1.00	0.00	0.949,1.000	0.000,0.051	R	R	9
3100H-99293	0	1.00	0.00	0.957,1.000	0.000,0.043	R	R	9
3100H-99512	0	1.00	0.00	0.964,1.000	0.000,0.036	R	R	9
4100L-68624	0	1.00	0.00	0.978,1.000	0.000,0.022	R	R	11
4100L-68626	2	1.00	0.00	0.957,1.000	0.000,0.043	R	R	9
4100L-68631	2	1.00	0.00	0.955,1.000	0.000,0.045	R	R	9
4100L-68635	2	1.00	0.00	0.978,1.000	0.000,0.022	R	R	11
9000E-57680	0	1.00	0.00	0.957,1.000	0.000,0.043	R	R	9
9000E-57685	2	1.00	0.00	0.976,1.000	0.000,0.024	R	R	11
E44516	0	0.63	0.37	0.285,0.907	0.093,0.715	B	Uncertain	1
H25244	**0**	**0.00**	**1.00**	**0.000,0.031**	**0.969,1.000**	**R**	**B**	**5**
H25245	5	1.00	0.00	0.973,1.000	0.000,0.027	R	R	16
H25246	0	1.00	0.00	0.976,1.000	0.000,0.024	R	R	17
H25247	0	1.00	0.00	0.956,1.000	0.000,0.044	R	R	9
H25248	19	0.13	0.87	0.000,1.000	0.000,1.000	R	Uncertain	18
H25249	0	0.45	0.55	0.192,0.728	0.272,0.808	B	Uncertain	19
H25250	22	0.45	0.55	0.000,1.000	0.000,1.000	B	Uncertain	20
H25251	**0**	**1.00**	**0.00**	**0.977,1.000**	**0.000,0.023**	**B**	**R**	**11**

Appendix B: Continued

Band	% Mis.	Inferred Clusters		95% Probability Intervals		Field ID	Structure ID	Haplotype
H25252	**2**	**1.00**	**0.00**	**0.951,1.000**	**0.000,0.049**	**B**	**R**	**9**
H25256	**11**	**1.00**	**0.00**	**0.951,1.000**	**0.000,0.049**	**B**	**R**	**21**
H25257	**0**	**1.00**	**0.00**	**0.957,1.000**	**0.000,0.043**	**B**	**R**	**9**
H25259	**11**	**1.00**	**0.00**	**0.952,1.000**	**0.000,0.048**	**B**	**R**	**22**
H25260	**0**	**1.00**	**0.00**	**0.975,1.000**	**0.000,0.025**	**B**	**R**	**23**
H25330	**11**	**1.00**	**0.00**	**0.946,1.000**	**0.000,0.055**	**B**	**R**	**24**
H25331	8	0.44	0.57	0.140,0.770	0.230,0.860	B	Uncertain	25
H25334	2	0.77	0.23	0.425,1.000	0.000,0.575	B	Uncertain	26
H25335	5	1.00	0.00	0.971,1.000	0.000,0.029	R	R	27
H25336	11	0.64	0.36	0.355,0.886	0.114,0.645	B	Uncertain	28
H25337	0	0.63	0.37	0.284,0.909	0.091,0.716	B	Uncertain	29
H25338	**2**	**1.00**	**0.00**	**0.976,1.000**	**0.000,0.024**	**B**	**R**	**30**
H25339	**13**	**1.00**	**0.01**	**0.944,1.000**	**0.000,0.056**	**B**	**R**	**31**
H25340	0	0.00	1.00	0.000,0.032	0.968,1.000	B	B	5
H25341	**16**	**0.82**	**0.18**	**0.526,1.000**	**0.000,0.474**	**B**	**R**	**32**
H25342	16	0.66	0.34	0.298,1.000	0.000,0.702	B	Uncertain	33
H25343	**5**	**1.00**	**0.00**	**0.976,1.000**	**0.000,0.024**	**B**	**R**	**11**
H25345	2	1.00	0.00	0.962,1.000	0.000,0.038	R	R	9
H25346	0	0.00	1.00	0.000,0.033	0.967,1.000	B	B	5
H25347	**0**	**1.00**	**0.00**	**0.976,1.000**	**0.000,0.024**	**B**	**R**	**11**
H25348	**2**	**1.00**	**0.00**	**0.964,1.000**	**0.000,0.036**	**B**	**R**	**34**
H25349	**2**	**1.00**	**0.00**	**0.955,1.000**	**0.000,0.045**	**B**	**R**	**9**
H25350	**0**	**1.00**	**0.00**	**0.978,1.000**	**0.000,0.022**	**B**	**R**	**11**
H25351	0	0.00	1.00	0.000,0.029	0.971,1.000	B	B	5
H25352	0	0.00	1.00	0.000,0.033	0.967,1.000	B	B	5
H25353	0	0.00	1.00	0.000,0.030	0.970,1.000	B	B	5
H25354	0	0.00	1.00	0.000,0.029	0.971,1.000	B	B	5
H25355	0	0.00	1.00	0.000,0.031	0.969,1.000	B	B	5

Appendix B: Continued

Band	% Mis.	Inferred Clusters		95% Probability Intervals		Field ID	Structure ID	Haplotype
H25356	**0**	**1.00**	**0.00**	**0.975,1.000**	**0.000,0.025**	**B**	**R**	**11**
H25357	5	0.00	1.00	0.000,0.030	0.970,1.000	B	B	5
H25358	**8**	**1.00**	**0.01**	**0.941,1.000**	**0.000,0.059**	**B**	**R**	**35**
H25359	0	0.00	1.00	0.000,0.032	0.968,1.000	B	B	5
H25360	0	0.00	1.00	0.000,0.027	0.973,1.000	B	B	10
H25361	0	0.00	1.00	0.000,0.039	0.961,1.000	B	B	5
H25362	5	0.00	1.00	0.000,0.029	0.971,1.000	B	B	36
H25363	0	0.00	1.00	0.000,0.029	0.971,1.000	B	B	5
H25364	0	0.00	1.00	0.000,0.031	0.969,1.000	B	B	5
H25365	0	0.00	1.00	0.000,0.026	0.974,1.000	B	B	5
H25366	0	0.00	1.00	0.000,0.027	0.973,1.000	B	B	5
H25368	0	0.00	1.00	0.000,0.031	0.969,1.000	B	B	5
H25369	11	0.01	1.00	0.000,0.066	0.934,1.000	B	B	37
H25370	2	0.00	1.00	0.000,0.032	0.968,1.000	B	B	38
H25371	0	0.00	1.00	0.000,0.033	0.967,1.000	B	B	2
H25372	2	0.00	1.00	0.000,0.044	0.956,1.000	B	B	39
H25373	2	0.00	1.00	0.000,0.034	0.966,1.000	B	B	4
H25374	0	0.00	1.00	0.000,0.029	0.971,1.000	B	B	5
H25375	0	0.00	1.00	0.000,0.030	0.970,1.000	B	B	5
H25376	13	0.66	0.34	0.374,0.890	0.110,0.626	B	Uncertain	40
H25379	2	0.00	1.00	0.000,0.032	0.969,1.000	B	B	5
H25380	2	0.00	1.00	0.000,0.028	0.972,1.000	B	B	5
H25381	0	0.00	1.00	0.000,0.040	0.960,1.000	B	B	5
H25382	**16**	**0.99**	**0.01**	**0.892,1.000**	**0.000,0.108**	**B**	**R**	**41**
H25383	19	0.21	0.79	0.000,1.000	0.000,1.000	B	Uncertain	42
H25384	2	0.00	1.00	0.000,0.028	0.972,1.000	B	B	10
H25385	2	0.00	1.00	0.000,0.035	0.965,1.000	B	B	4
H25386	0	0.00	1.00	0.000,0.032	0.968,1.000	B	B	5

Appendix B: Continued

Band	% Mis.	Inferred Clusters		95% Probability Intervals		Field ID	Structure ID	Haplotype
H25387	11	0.00	1.00	0.000,0.046	0.954,1.000	B	B	43
H25388	0	0.00	1.00	0.000,0.031	0.969,1.000	B	B	5
H25389	0	0.00	1.00	0.000,0.029	0.971,1.000	B	B	5
H25390	0	0.00	1.00	0.000,0.032	0.968,1.000	B	B	5
H25391	16	0.16	0.84	0.000,1.000	0.000,1.000	B	Uncertain	44
H25392	0	0.00	1.00	0.000,0.028	0.972,1.000	B	B	5
H25393	**13**	**1.00**	**0.00**	**0.963,1.000**	**0.000,0.037**	**B**	**R**	**45**
H25394	0	0.00	1.00	0.000,0.030	0.970,1.000	B	B	5
H25395	**0**	**1.00**	**0.00**	**0.977,1.000**	**0.000,0.023**	**B**	**R**	**11**
H25396	0	1.00	0.00	0.957,1.000	0.000,0.043	R	R	7
H25397	**0**	**0.00**	**1.00**	**0.000,0.029**	**0.971,1.000**	**R**	**B**	**5**
H25398	16	0.96	0.04	0.497,1.000	0.000,0.503	R	Uncertain	46
H25399	**8**	**0.00**	**1.00**	**0.000,0.031**	**0.969,1.000**	**R**	**B**	**5**
H25400	5	1.00	0.00	0.962,1.000	0.000,0.038	R	R	47
L34317	**0**	**1.00**	**0.00**	**0.970,1.000**	**0.000,0.030**	**B**	**R**	**11**
L34341	**16**	**0.99**	**0.01**	**0.896,1.000**	**0.000,0.104**	**B**	**R**	**48**
L34342	25	0.45	0.55	0.000,1.000	0.000,1.000	B	Uncertain	49
L34344	0	0.00	1.00	0.000,0.031	0.970,1.000	B	B	5
L34350	0	0.00	1.00	0.000,0.031	0.969,1.000	B	B	2
L34351	0	0.00	1.00	0.000,0.031	0.969,1.000	B	B	5
L34403	0	0.45	0.55	0.194,0.731	0.269,0.806	B	Uncertain	50
L34404	0	0.00	1.00	0.000,0.030	0.970,1.000	B	B	5
L34405	0	0.00	1.00	0.000,0.030	0.970,1.000	B	B	5
L34410	2	0.00	1.00	0.000,0.028	0.972,1.000	B	B	51
L34414	16	0.06	0.94	0.000,1.000	0.000,1.000	B	Uncertain	52
L34415	0	0.00	1.00	0.000,0.026	0.974,1.000	B	B	5
L34910	27	0.45	0.55	0.000,1.000	0.000,1.000	B	Uncertain	53
L34922	**2**	**1.00**	**0.00**	**0.977,1.000**	**0.000,0.023**	**B**	**R**	**8**

Appendix B: Continued

Band	% Mis.	Inferred Clusters		95% Probability Intervals		Field ID	Structure ID	Haplotype
L35651	0	0.00	1.00	0.000,0.027	0.973,1.000	B	B	5
L35652	0	0.00	1.00	0.000,0.035	0.965,1.000	B	B	5
L35653	0	0.00	1.00	0.000,0.032	0.968,1.000	B	B	5
L35654	**22**	**1.00**	**0.00**	**0.956,1.000**	**0.000,0.044**	**B**	**R**	**54**
L4100-68539	0	0.00	1.00	0.000,0.027	0.973,1.000	B	B	55
L4100-68541	0	0.00	1.00	0.000,0.025	0.975,1.000	B	B	6
L4100-68542	**0**	**1.00**	**0.00**	**0.976,1.000**	**0.000,0.024**	**B**	**R**	**11**
L4100-68543	0	0.00	1.00	0.000,0.028	0.972,1.000	B	B	5
L4100-68544	**0**	**1.00**	**0.00**	**0.974,1.000**	**0.000,0.026**	**B**	**R**	**11**
L4100-68545	0	0.00	1.00	0.000,0.028	0.972,1.000	B	B	5
L4100-68546	0	0.00	1.00	0.000,0.029	0.971,1.000	B	B	5
L4100-68549	0	0.00	1.00	0.000,0.027	0.973,1.000	B	B	5
L4100-68550	0	0.00	1.00	0.000,0.029	0.971,1.000	B	B	5
L4100-68551	**0**	**1.00**	**0.00**	**0.975,1.000**	**0.000,0.025**	**B**	**R**	**56**
L60380	**0**	**1.00**	**0.00**	**0.953,1.000**	**0.000,0.047**	**B**	**R**	**9**
L60382	**2**	**1.00**	**0.00**	**0.975,1.000**	**0.000,0.025**	**B**	**R**	**8**
L60383	**2**	**1.00**	**0.00**	**0.958,1.000**	**0.000,0.042**	**B**	**R**	**57**
L60384	2	1.00	0.00	0.959,1.000	0.000,0.041	U	R	58
L60386	0	0.00	1.00	0.000,0.029	0.971,1.000	B	B	5
L60387	**0**	**1.00**	**0.00**	**0.975,1.000**	**0.000,0.025**	**B**	**R**	**11**
L60388	**2**	**1.00**	**0.00**	**0.962,1.000**	**0.000,0.038**	**B**	**R**	**59**
L60389	**2**	**1.00**	**0.00**	**0.956,1.000**	**0.000,0.044**	**B**	**R**	**3**
L60391	**2**	**1.00**	**0.00**	**0.975,1.000**	**0.000,0.025**	**B**	**R**	**8**
L60392	0	0.00	1.00	0.000,0.028	0.972,1.000	B	B	5
L60393	0	0.00	1.00	0.000,0.032	0.968,1.000	B	B	60
L60394	**2**	**1.00**	**0.00**	**0.975,1.000**	**0.000,0.025**	**B**	**R**	**8**
L60395	**27**	**1.00**	**0.00**	**0.967,1.000**	**0.000,0.033**	**B**	**R**	**61**
L60396	**2**	**1.00**	**0.00**	**0.978,1.000**	**0.000,0.022**	**B**	**R**	**8**

Appendix B: Continued

Band	% Mis.	Inferred Clusters		95% Probability Intervals		Field ID	Structure ID	Haplotype
L60397	**2**	**1.00**	**0.00**	**0.974,1.000**	**0.000,0.026**	**B**	**R**	**8**
L60398	**2**	**1.00**	**0.00**	**0.972,1.000**	**0.000,0.028**	**B**	**R**	**8**
L60399	**16**	**0.99**	**0.01**	**0.825,1.000**	**0.000,0.175**	**B**	**R**	**62**
L60401	**0**	**1.00**	**0.00**	**0.946,1.000**	**0.000,0.054**	**B**	**R**	**63**
L60403	**0**	**1.00**	**0.00**	**0.962,1.000**	**0.000,0.038**	**B**	**R**	**9**
L60404	**0**	**1.00**	**0.00**	**0.962,1.000**	**0.000,0.038**	**B**	**R**	**64**
L60405	**2**	**1.00**	**0.00**	**0.968,1.000**	**0.000,0.032**	**B**	**R**	**8**
L60406	0	0.00	1.00	0.000,0.028	0.972,1.000	B	B	5
L60407	**0**	**1.00**	**0.00**	**0.953,1.000**	**0.000,0.047**	**B**	**R**	**9**
L60408	**0**	**1.00**	**0.00**	**0.974,1.000**	**0.000,0.026**	**B**	**R**	**11**
L60409	2	0.00	1.00	0.000,0.030	0.970,1.000	B	B	65
L60410	**0**	**1.00**	**0.00**	**0.957,1.000**	**0.000,0.043**	**B**	**R**	**7**
L60411	0	0.00	1.00	0.000,0.031	0.969,1.000	B	B	5
L60412	25	0.45	0.55	0.000,1.000	0.000,1.000	B	Uncertain	66
L60451	0	1.00	0.00	0.978,1.000	0.000,0.022	R	R	11
L60452	0	1.00	0.00	0.976,1.000	0.000,0.024	R	R	11
L60454	2	1.00	0.00	0.971,1.000	0.000,0.029	R	R	67
L60456	0	1.00	0.00	0.958,1.000	0.000,0.042	R	R	9
L60457	0	1.00	0.00	0.963,1.000	0.000,0.037	R	R	68
L60458	0	1.00	0.00	0.960,1.000	0.000,0.040	R	R	9
L60459	0	1.00	0.00	0.956,1.000	0.000,0.044	R	R	9
L60460	0	1.00	0.00	0.977,1.000	0.000,0.023	R	R	11
L60461	0	1.00	0.00	0.977,1.000	0.000,0.023	R	R	11
L60462	2	1.00	0.00	0.954,1.000	0.000,0.046	R	R	3
L60463	2	0.83	0.17	0.533,1.000	0.000,0.467	R	R	69
L60464	0	1.00	0.00	0.975,1.000	0.000,0.025	R	R	11
L60465	5	1.00	0.00	0.956,1.000	0.000,0.044	R	R	3
L60466	0	1.00	0.00	0.962,1.000	0.000,0.038	R	R	9

Appendix B: Continued

Band	% Mis.	Inferred Clusters		95% Probability Intervals		Field ID	Structure ID	Haplotype
L60467	0	1.00	0.00	0.958,1.000	0.000,0.042	R	R	9
L60468	2	1.00	0.00	0.955,1.000	0.000,0.045	R	R	3
L60469	0	1.00	0.00	0.976,1.000	0.000,0.024	R	R	11
L60470	0	1.00	0.00	0.965,1.000	0.000,0.035	R	R	9
L60471	0	1.00	0.00	0.953,1.000	0.000,0.047	R	R	9
L60472	0	0.00	1.00	0.000,0.030	0.970,1.000	B	B	5
L60473	0	0.00	1.00	0.000,0.029	0.971,1.000	B	B	6
L60474	**0**	**1.00**	**0.00**	**0.978,1.000**	**0.000,0.022**	**B**	**R**	**70**
L60475	**0**	**1.00**	**0.00**	**0.975,1.000**	**0.000,0.025**	**B**	**R**	**71**
L60476	**2**	**1.00**	**0.00**	**0.976,1.000**	**0.000,0.024**	**B**	**R**	**8**
L60477	2	1.00	0.00	0.952,1.000	0.000,0.048	R	R	9
L60478	0	1.00	0.00	0.975,1.000	0.000,0.025	R	R	11
L60479	0	1.00	0.00	0.975,1.000	0.000,0.025	R	R	11
L60481	0	1.00	0.00	0.976,1.000	0.000,0.024	R	R	11
L60482	0	1.00	0.00	0.974,1.000	0.000,0.026	R	R	11
L60483	0	1.00	0.00	0.957,1.000	0.000,0.043	R	R	9
L60484	0	1.00	0.00	0.956,1.000	0.000,0.044	R	R	9
L60485	0	1.00	0.00	0.962,1.000	0.000,0.038	R	R	9
L60486	0	1.00	0.00	0.957,1.000	0.000,0.043	R	R	9
L60487	0	1.00	0.00	0.975,1.000	0.000,0.025	R	R	11
L60488	0	1.00	0.00	0.969,1.000	0.000,0.031	R	R	11
L60489	0	1.00	0.00	0.960,1.000	0.000,0.040	R	R	9
L60490	0	1.00	0.00	0.957,1.000	0.000,0.043	R	R	9
L60491	0	1.00	0.00	0.956,1.000	0.000,0.044	R	R	9
L60492	**0**	**0.00**	**1.00**	**0.000,0.036**	**0.964,1.000**	**R**	**B**	**5**
L60493	2	1.00	0.00	0.956,1.000	0.000,0.044	R	R	9
L60494	0	1.00	0.00	0.954,1.000	0.000,0.046	R	R	9
L60495	0	1.00	0.00	0.962,1.000	0.000,0.038	R	R	9

LITERATURE CITED

Adhikari, D., A. Chettri, and S. K. Barik. 2009. Modelling the ecology and distribution of highly pathogenic avian influenza (H5N1) in the Indian subcontinent. Current Science (Bangalore) 97:73–78.

Allouche, O., O. Steinitz, D. Rotem, A. Rosenfeld, and R. Kadmon. 2008. Incorporating distance constraints into species distribution models. Journal of Applied Ecology 45:599–609.

Anderson, B. W., and R. J. Daugherty. 1974. Characteristics and reproductive biology of Grosbeaks (*Pheucticus*) in the hybrid zone in South Dakota. Wilson Bulletin 86:1–11.

Anderson, R. P., M. Gómez-Laverde, and A. T. Peterson. 2002. Geographical distributions of Spiny Pocket Mice in South America: Insights from predictive models. Global Ecology and Biogeography 11:131–141.

Anderson, R. P., D. Lew, and A. T. Peterson. 2003. Evaluating predictive models of species' distributions: criteria for selecting optimal models. Ecological modelling 162:211–232.

Avise, J. C., and K. Wollenberg. 1997. Phylogenetics and the origin of species. Proceedings of the National Academy of Sciences 94:7748–7755.

Baker, R. J., and R. D. Bradley. 2006. Speciation in mammals and the genetic species concept. Journal of Mammalogy 87:643–662.

Baker, R. J., S. K. Davis, R. D. Bradley, M. J. Hamilton, and R. A. van D. Bussche. 1989. Ribosomal-DNA, mitochondrial-DNA, chromosomal, and allozymic studies on a contact zone in the Pocket Gopher, *Geomys*. Evolution 43:63–75.

Baker, R. J., C. J. Phillips, R. D. Bradley, J. M. Burns, D. Cooke, G. F. Edson, D. R. Haragan, C. Jones, R. R. Monk, J. T. Montford, D. J. Schmidly, and N. C. Parker. 1998.

Bioinformative, museums and society: Integrating biological data for knowledge-based decisions. Occasional papers museum of Texas Tech University 187.

Baltosser, W. H., and S. M. Russell. 2000. Black-chinned Hummingbird: *Archilochus alexandri*. Birds of North America. Birds of North America, Inc., Philadelphia, PA, USA.

Barton, N. 1979. The dynamics of hybrid zones. Heredity 43:341.

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

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

Bass, M. S., M. Finer, C. N. Jenkins, H. Kreft, D. F. Cisneros-Heredia, S. F. McCracken, N. C. A. Pitman, P. H. English, K. Swing, G. Villa, A. Di Fiore, C. C. Voigt, and T. H. Kunz. 2010. Global conservation significance of Ecuador's Yasuní National Park. PLoS ONE 5:e8767.

Bennett, A. F., R. E. Lenski, and J. E. Mittler. 1992. Evolutionary adaptation to temperature. I. Fitness responses of *Escherichia coli* to changes in its thermal environment. Evolution 46:16–30.

Bergen, K. M., A. M. Gilboy, and D. G. Brown. 2007. Multi-dimensional vegetation structure in modeling avian habitat. Ecological Informatics 2:9–22.

BirdLife International and NatureServe. 2012. Bird species distribution maps of the world. BirdLife International, Cambridge, UK and NatureServe, Arlington, TX, USA.

Bleiweiss, R. 1998. Tempo and mode of hummingbird evolution. Biological Journal of the Linnean Society 65:63–76.

Braunisch, V., K. Bollmann, R. F. Graf, and A. H. Hirzel. 2008. Living on the edge—Modelling habitat suitability for species at the edge of their fundamental niche. Ecological Modelling 214:153–167.

Brown, J. H., and M. V. Lomolino. 1998. Biogeography. 2nd edition. Sinauer Associates, Sunderland, MA, USA.

Brunsfeld, S. J., J. Sullivan, D. E. Soltis, and P. S. Soltis. 2001. Comparative phylogeography of northwestern North America: a synthesis. Pages 319–340 *in* J. Silvertown and J. Antonovics, editors. Integrating Ecology and Evolution in a Spatial Context. Blackwell Science, Oxford, UK.

Bush, K. L., M. D. Vinsky, C. L. Aldridge, and C. A. Paszkowski. 2005. A comparison of sample types varying in invasiveness for use in DNA sex determination in an endangered population of Greater Sage-Grouse (*Centrocercus uropihasianus*). Conservation Genetics 6:867–870.

Butler, C. J., J. L. Curtis, K. McBride, D. Arbour, and B. Heck. 2011. Modeling the distribution of the Dwarf Palmetto (*Sabal minor*; Arecaceae) in McCurtain County, Oklahoma. The Southwestern Naturalist 56:66–70.

Butler, C. J., D. Ledbetter, N. Batchelder, and G. Batchelder. 2007. Black-chinned Hummingbirds breed in Grady County. Bulletin of the Oklahoma Ornithological Society 40:13–16.

Butler, C. J., E. A. Wheeler, and L. B. Stabler. 2012. Distribution of the threatened Lace Hedgehog Cactus (*Echinocereus reichenbachii*) under various climate change scenarios. The Journal of the Torrey Botanical Society 139:46–55.

Cardinale, B. J. 2011. Biodiversity improves water quality through niche partitioning. Nature 472:86–89.

Chatfield, M. W. H., K. H. Kozak, B. M. Fitzpatrick, and P. K. Tucker. 2010. Patterns of differential introgression in a salamander hybrid zone: inferences from genetic data and ecological niche modelling. Molecular Ecology 19:4265–4282.

Chen, G., and T. A. Peterson. 1999. A new technique for predicting distribution of terrestrial vertebrates using inferential modeling. Zoological Research 21:231–237.

Chen, Y., and J. Bi. 2007. Biogeography and hotspots of amphibian species of China: implications to reserve selection and conservation. Current Science 92:480–489.

Chubb, A. L. 2004. Nuclear corroboration of DNA–DNA hybridization in deep phylogenies of hummingbirds, swifts, and passerines: the phylogenetic utility of ZENK (ii). Molecular Phylogenetics and Evolution 30:128–139.

Cicero, C. 2004. Barriers to sympatry between avian sibling species (Paridae: *Baeolophus*) in local secondary contact. Evolution 58:1573–1587.

Constable, H., R. Guralnick, J. Wieczorek, C. Spencer, A. T. Peterson, and The VertNet Steering Committee. 2010. VertNet: A new model for biodiversity data sharing. PLoS Biology 8:e1000309.

Coyne, J. A., and H. A. Orr. 2004. Speciation. W.H. Freeman, New York, NY, USA.

Cracraft, J. 1983. Species concepts and speciation analysis. Pages 159–87 *in* R. F. Johnston, editor. Current Ornithology. Plenum Press, New York, NY, USA and London, UK.

Crandall, K. A., A. R. Templeton, and C. F. Sing. 1994. Intraspecific phylogenetics: problems and solutions. Pages 273–297 *in* R. W. Scotland, D. J. Siebert, and D. M. Williams,

editors. Models in Phylogeny Reconstruction. The Systematics Association Special Volume No 52, Clarendon Press, Oxford, UK.

Culumber, Z. W., D. B. Shepard, S. W. Coleman, G. G. Rosenthal, and M. Tobler. 2012. Physiological adaptation along environmental gradients and replicated hybrid zone structure in swordtails (Teleostei: *Xiphophorus*). Journal of Evolutionary Biology 25:1800–1814.

Curry, C. M., and M. A. Patten. 2014. Current and historical extent of phenotypic variation in the Tufted and Black-crested Titmouse (Paridae) hybrid zone in the southern Great Plains. The American Midland Naturalist 171:271–300.

Donoghue, M. J. 1985. A critique of the biological species concept and recommendations for a phylogenetic alternative. The Bryologist 88:172–181.

Earl, D. A., and B. M. vonHoldt. 2011. Structure Harvester: a website and program for visualizing Structure output and implementing the Evanno method. Conservation Genetics Resources 4:359–361.

Edgar, R. C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Research 32:1792–1797.

Elith, J., C. H. Graham, R. P. Anderson, M. Dudík, S. Ferrier, A. Guisan, R. J. Hijmans, F. Huettmann, J. R. Leathwick, A. Lehmann, J. Li, L. G. Lohmann, B. A. Loiselle, G. Manion, C. Moritz, M. Nakamura, Y. Nakazawa, J. McC. M. Overton, A. Townsend Peterson, S. J. Phillips, K. Richardson, R. Scachetti-Pereira, R. E. Schapire, J. Soberón, S. Williams, M. S. Wisz, and N. E. Zimmermann. 2006. Novel methods improve prediction of species' distributions from occurrence data. Ecography 29:129–151.

Elith, J., S. J. Phillips, T. Hastie, M. Dudík, Y. E. Chee, and C. J. Yates. 2011. A statistical explanation of MaxEnt for ecologists. Diversity and Distributions 17:43–57.

Emlen, S. T., J. D. Rising, and W. L. Thompson. 1975. A behavioral and morphological study of sympatry in the Indigo and Lazuli Buntings of the Great Plains. The Wilson Bulletin 87:145–179.

Endler, J. A. 1977. Geographic Variation, Speciation, and Clines. Princeton University Press, Princeton, NJ, USA.

Estrada-Peña, A., and J. M. Venzal. 2007. Climate niches of tick species in the Mediterranean region: Modeling of occurrence data, distributional constraints, and impact of climate change. Journal of Medical Entomology 44:1130–1138.

Felsenstein, J. 1985. Confidence limits on phylogenies: An approach using the bootstrap. Evolution 39:783–791.

Ficken, M. S., and R. W. Ficken. 1968. Territorial relationships of Blue-winged Warblers, Golden-winged Warblers, and their hybrids. The Wilson Bulletin 80:442–451.

Funk, V. A., M. F. Zermoglio, and N. Nasir. 1999. Testing the use of specimen collection data and GIS in biodiversity exploration and conservation decision making in Guyana. Biodiversity and Conservation 8:727–751.

Gallego, D., F. Cánovas, M. A. Esteve, and J. Galián. 2004. Descriptive biogeography of *Tomicus* (Coleoptera: Scolytidae) species in Spain. Journal of Biogeography 31:2011–2024.

García, A. 2006. Using ecological niche modelling to identify diversity hotspots for the herpetofauna of Pacific lowlands and adjacent interior valleys of Mexico. Biological Conservation 130:25–46.

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

Gill, F. B. 2007. Ornithology. 3rd edition. W.H. Freeman and Co., New York, NY, USA.

Goudet, J. 2001. FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). Available from http://www.unil.ch/izea/softwares/fstat.html. Updated from Goudet (1995).

Graham, A. 1999. Late Cretaceous and Cenozoic History of North American Vegetation, North of Mexico. Oxford University Press, UK.

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

Green, R. H. 1971. A multivariate statistical approach to the Hutchinsonian niche: Bivalve molluscs of central Canada. Ecology 52:544–556.

Grinnell, J. 1917. The niche-relationships of the California Thrasher. The Auk 34:427–433.

Guralnick, R., and H. Constable. 2010. VertNet: Creating a data-sharing community. BioScience 60:258–259.

Gustafson, M. E., J. Hildenbrand, and L. Metras. 1997. The North American bird banding manual (Electronic Version). Version 1.0. Courtesy of the U.S. Geological Survey.

Hardy, G. H. 1908. Mendelian proportions in a mixed population. Science 28:49–50.

Harrison, R. G. 1993. Hybrid Zones and the Evolutionary Process. Oxford University Press, UK.

Hebert, P. D. N., S. Ratnasingham, and J. R. de Waard. 2003. Barcoding animal life: cytochrome c oxidase subunit I divergences among closely related species. Proceedings of the Royal Society of London B: Biological Sciences 270:S96–S99.

Hebert, P. D. N., M. Y. Stoeckle, T. S. Zemlak, and C. M. Francis. 2004. Identification of birds through DNA barcodes. PLoS Biology 2:e312.

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

Hijmans, R. J., S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis. 2005. Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology 25:1965–1978.

Hirzel, A. H., and R. Arlettaz. 2003. Modeling habitat suitability for complex species distributions by environmental-distance geometric mean. Environmental Management 32:614–623.

Hirzel, A. H., J. Hausser, D. Chessel, and N. Perrin. 2002a. Ecological-niche factor analysis: how to compute habitat-suitability maps without absence data? Ecology 83:2027–2036.

Hirzel, A., J. Hausser, and N. Perrin. 2002b. Biomapper 3.1. Labaoratory for Conservation Biology, Lausanne, Switzerland.

Hirzel, A. H., G. Le Lay, V. Helfer, C. Randin, and A. Guisan. 2006. Evaluating the ability of habitat suitability models to predict species presences. Ecological Modelling 199:142–152.

Holdridge, L. R. 1945. A brief sketch of the flora of Hispaniola. Pages 76–78 *in* F. Verdoom, editor. Plants and Plant Science in Latin America. Chronica Botanica Company, Waltham, MA, USA.

Howell, S. N. G. 2002. Hummingbirds of North America: The Photographic Guide. Academic Press, San Diego, CA, USA.

Hutchinson, G. E. 1957. Concluding remarks. Cold Spring Harbor symposia on quantitative biology 22:415–427.

Jaynes, E. T. 1957. Information theory and statistical mechanics. Physical Review 106:620–630.

Johnson, N. K., and C. Cicero. 2004. New mitochondrial DNA data affirm the importance of Pleistocene speciation in North American birds. Evolution 58:1122–1130.

Judd, E. R., C. J. Butler, and N. Batchelder. 2011. Hybridization between Black-chinned (*Archilochus alexandri*) and Ruby-throated (*A. colubris*) Hummingbirds in Oklahoma. Bulletin of the Oklahoma Ornithological Society 44:1–7.

Kimura, M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. Journal of Molecular Evolution 16:111–120.

Kohlmann, B., H. Nix, and D. D. Shaw. 1988. Environmental predictions and distributional limits of chromosomal taxa in the Australian grasshopper *Caledia captiva* (F.). Oecologia 75:483–493.

Kroodsma, R. L. 1974a. Species-recognition behavior of territorial male Rose-breasted and Black-headed Grosbeaks (*Pheucticus*). The Auk 91:54–64.

Kroodsma, R. L. 1974b. Hybridization in grosbeaks (*Pheucticus*) in North Dakota. The Wilson Bulletin 86:230–236.

Kroodsma, R. L. 1975. Hybridization in buntings (*Passerina*) in North Dakota and eastern Montana. The Auk 92:66–80.

Lance, S. L., C. Hagen, T. C. Glenn, R. T. Brumfield, K. F. Stryjewski, and G. R. Graves. 2008. Fifteen polymorphic microsatellite loci from Jamaican Streamertail Hummingbirds (*Trochilus*). Conservation Genetics 10:1195–1198.

Lawler, S. P., and P. J. Morin. 1993. Temporal overlap, competition, and priority effects in larval anurans. Ecology 74:174.

Martínez-Freiría, F., N. Sillero, M. Lizana, and J. C. Brito. 2008. GIS-based niche models identify environmental correlates sustaining a contact zone between three species of European vipers. Diversity and Distributions 14:452–461.

Mayr, E. 1970. Populations, Species, and Evolution; An Abridgment of Animal Species and Evolution. Belknap Press of Harvard University Press, Cambridge, MA, USA.

McGuire, J. A., C. C. Witt, J. V. R. Jr, R. Dudley, and D. L. Altshuler. 2009. A higher-level taxonomy for hummingbirds. Journal of Ornithology 150:155–165.

McGuire, J. A., C. C. Witt, J. V. Remsen, A. Corl, D. L. Rabosky, D. L. Altshuler, and R. Dudley. 2014. Molecular phylogenetics and the diversification of hummingbirds. Current Biology 24:910–916.

McGuire, J., C. Witt, D. Altshuler, and J. V. Remsen. 2007. Phylogenetic systematics and biogeography of hummingbirds: Bayesian and maximum likelihood analyses of partitioned data and selection of an appropriate partitioning strategy. Systematic Biology 56:837–856.

McKitrick, M. C., and R. M. Zink. 1988. Species concepts in ornithology. The Condor 90:1–14.

Merriam, C. H. 1898. Life zones and crop zones of the United States. US Government Printing Office, Washington, D.C., USA.

Miraldo, A., C. Faria, G. M. Hewitt, O. S. Paulo, and B. C. Emerson. 2013. Genetic analysis of a contact zone between two lineages of the Ocellated Lizard (*Lacerta lepida;* Daudin 1802) in south-eastern Iberia reveal a steep and narrow hybrid zone. Journal of Zoological Systematics and Evolutionary Research 51:45–54.

Moore, W. S. 1977. An evaluation of narrow hybrid zones in vertebrates. The Quarterly Review of Biology 52:263.

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

Moore, W. S., J. H. Graham, and J. T. Price. 1991. Mitochondrial DNA variation in the Northern Flicker (*Colaptes auratus*, Aves). Molecular Biology and Evolution 8:327–344.

Moore, W. S., and W. D. Koenig. 1986. Comparative reproductive success of yellow-shafted, red-shafted, and hybrid flickers across a hybrid zone. The Auk: 42–51.

Moore, W. S., and J. T. Price. 1993. Nature and selection in the Northern Flicker hybrid zone and its implications for speciation theory. Pages 196–225 *in* R. G. Harrison, editor. Hybrid Zones and the Evolutionary Process. Oxford University Press US, New York, NY, USA.

Muñoz, M. E. S., R. Giovanni, M. F. Siqueira, T. Sutton, P. Brewer, R. S. Pereira, D. A. L. Canhos, and V. P. Canhos. 2009. openModeller: A generic approach to species' potential distribution modelling. GeoInformatica 15:111–135.

Nei, M. 1987. Molecular Evolutionary Genetics. Columbia University Press, New York, NY, USA.

Nei, M., T. Maruyama, and R. Chakraborty. 1975. The bottleneck effect and genetic variability in populations. Evolution 29:1–10.

Nix, H. 1986. A biogeographic analysis of Australian elapid snakes. Pages 4–16 *in* R. Longmore, editor. Atlas of Elapid Snakes of Australia. Bureau of Flora and Fauna, Canberra, Australia.

Nixon, K. C., and Q. D. Wheeler. 1990. An amplification of the phylogenetic species concept. Cladistics 6:211–223.

Odum, E. P. 1969. The strategy of ecosystem development. Science 164:262–270.

Olson, R. A., J. D. Winter, D. C. Nettles, and J. M. Haynes. 1988. Resource partitioning in summer by salmonids in south-central Lake Ontario. Transactions of the American Fisheries Society 117:552–559.

Orr, H. A. 2005. The genetic theory of adaptation: A brief history. Nature Reviews Genetics 6:119–127.

Papeş, M., and P. Gaubert. 2007. Modelling ecological niches from low numbers of occurrences: Assessment of the conservation status of poorly known viverrids (Mammalia, Carnivora) across two continents: Ecological niche modelling of poorly known viverrids. Diversity and Distributions 13:890–902.

Parker, D. T. 1987. Evolutionary genetics of House Sparrows. Pages 381–406 *in* F. Cooke and P. A. Buckley, editors. Avian Genetics: A Population and Ecological Approach. Academic Press, New York, NY, USA.

Peterson, A. T., M. Papeş, and M. Eaton. 2007. Transferability and model evaluation in ecological niche modeling: a comparison of GARP and Maxent. Ecography 30:550–560.

Peterson, A. T., and D. A. Vieglais. 2001. Predicting species invasions using ecological niche modeling: New approaches from bioinformatics attack a pressing problem. BioScience 51:363–371.

Petraitis, P. S. 1979. Likelihood measures of niche breadth and overlap. Ecology 60:703–710.

Phillips, S. J., R. P. Anderson, and R. E. Schapire. 2006. Maximum entropy modeling of species geographic distributions. Ecological Modelling 190:231–259.

Phillips, S. J., and M. Dudík. 2008. Modeling of species distributions with Maxent: New extensions and a comprehensive evaluation. Ecography 31:161–175.

Phillips, S. J., M. Dudík, and R. E. Schapire. 2004. A maximum entropy approach to species distribution modeling. Pages 655-622 *in* Proceedings of the twenty-first international conference on machine learning. ACM Press, New York, NY, USA.

Pianka, E. R. 1974. Niche overlap and diffuse competition. Proceedings of the National Academy of Sciences 71:2141–2145.

Pierotti, R., and C. A. Annett. 1993. Hybridization and male parental investment in birds. The Condor 95:670–679.

Ponder, W. F., G. A. Carter, P. Flemons, and R. R. Chapman. 2001. Evaluation of museum collection data for use in biodiversity assessment. Conservation Biology 15:648–657.

Posada, D., and K. A. Crandall. 2001. Selecting the best-fit model of nucleotide substitution. Systematic Biology 50:580–601.

Price, T. 1998. Sexual selection and natural selection in bird speciation. Philosophical Transactions of the Royal Society of London Series B: Biological Sciences 353:251–260.

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

Pritchard, J. K., W. Wen, and D. Falush. 2003. Documentation for STRUCTURE software: Version 2. Available from http://pritch.bsd.uchicago.edu.

Projecto-Garcia, J., C. Natarajan, H. Moriyama, R. E. Weber, A. Fago, Z. A. Cheviron, R. Dudley, J. A. McGuire, C. C. Witt, and J. F. Storz. 2013. Repeated elevational transitions in hemoglobin function during the evolution of Andean hummingbirds. Proceedings of the National Academy of Sciences of the United States of America 110:20669–20674.

Pulich, W. M. 1988. The Birds of North Central Texas. First Edition. Texas A&M University Press, College Station, TX, USA.

Pyke, G. H. 1982. Local geographic distributions of bumblebees near Crested Butte, Colorado: Competition and community structure. Ecology 63:555.

Pyle, P. 1997. Identification Guide to North American Birds, Part I: Columbidae to Ploceidae. Slate Creek Press, Bolinas, CA, USA.

QGIS Development Team. 2014. QGIS geographic information system. Open Source Geospatial Foundation Project. Available from http://qgis.osgeo.org.

Rainey, P. B., and M. Travisano. 1998. Adaptive radiation in a heterogeneous environment. Nature 394:69–72.

R Development Core Team. 2008. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available from http://www.R-project.org.

Reinking, D. L., editor. 2004. Oklahoma Breeding Bird Atlas. University of Oklahoma Press, Norman, OK, USA.

Rising, J. D. 1970. Morphological variation and evolution in some North American orioles. Systematic Biology 19:315–351.

Rising, J. D. 1983a. The Great Plains hybrid zones. Pages 131–157 *in* R. F. Johnston, editor. Current Ornithology. Plenum Press, New York, NY, USA.

Rising, J. D. 1983b. The progress of oriole hybridization in Kansas. The Auk: 885–897.

Rising, J. D. 1996. The stability of the oriole hybrid zone in western Kansas. The Condor 98:658–663.

Rivera, J. H. V., M. A. O. Huerta, and R. Guerrero. 2008. Analysis of the distribution of Orange-breasted Bunting (*Passerina leclancherii*): An endemic species of Mexico's Pacific slope. Ornitología Neotropical 19:265–274.

Robinson, T. R., R. R. Sargent, and M. B. Sargent. 1996. Ruby-throated Hummingbird (*Archilochus colubris*). *In* A. Poole and F. Gill, editors, The Birds of North America. The Birds of North America, Inc, Philadelphia, PA, USA.

Rohwer, S., E. Bermingham, and C. Wood. 2001. Plumage and mitochondrial DNA haplotype variation across a moving hybrid zone. Evolution 55:405–422.

Rohwer, S., and C. Wood. 1998. Three hybrid zones between Hermit and Townsend's Warblers in Washington and Oregon. The Auk 115:284–310.

Root, T. 1988. Environmental factors associated with avian distributional boundaries. Journal of Biogeography 15:489–505.

Roura-Pascual, N., L. Brotons, A. T. Peterson, and W. Thuiller. 2009. Consensual predictions of potential distributional areas for invasive species: A case study of Argentine ants in the Iberian Peninsula. Biological Invasions 11:1017–1031.

Roy, M. S., J. C. Torres-Mura, and F. Hertel. 1998. Evolution and history of hummingbirds (Aves: Trochilidae) from the Juan Fernandez Islands, Chile. Ibis 140:265–273.

Rubin, E. S., C. J. Stermer, W. M. Boyce, and S. G. Torres. 2009. Assessment of predictive habitat models for Bighorn Sheep in California's Peninsular Ranges. The Journal of Wildlife Management 73:859–869.

Russell, S. M., and R. O. Russell. 2001. The North American Banders' Manual for Banding Hummingbirds. North American Banding Council, Point Reyes, CA, USA.

Saitou, N., and M. Nei. 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. Molecular Biology and Evolution 4:406–425.

Sánchez-Flores, E. 2007. GARP modeling of natural and human factors affecting the potential distribution of the invasives *Schismus arabicus* and *Brassica tournefortii* in "El Pinacate y Gran Desierto de Altar" Biosphere Reserve. Ecological Modelling 204:457–474.

Santika, T., and M. F. Hutchinson. 2009. The effect of species response form on species distribution model prediction and inference. Ecological Modelling 220:2365–2379.

Santos, X., J. C. Brito, N. Sillero, J. M. Pleguezuelos, G. A. Llorente, S. Fahd, and X. Parellada. 2006. Inferring habitat-suitability areas with ecological modelling techniques and GIS: A contribution to assess the conservation status of *Vipera latastei*. Biological Conservation 130:416–425.

Sato, M., and K. Sato. 2012. Maternal inheritance of mitochondrial DNA. Autophagy 8:424–425.

Sattler, T., F. Bontadina, A. H. Hirzel, and R. Arlettaz. 2007. Ecological niche modelling of two cryptic bat species calls for a reassessment of their conservation status. Journal of Applied Ecology 44:1188–1199.

Seneviratne, S. S., D. P. L. Toews, A. Brelsford, and D. E. Irwin. 2012. Concordance of genetic and phenotypic characters across a sapsucker hybrid zone. Journal of Avian Biology 43:119–130.

Short, L. L. 1965. Hybridization in the flickers (*Colaptes*) of North America. Bulletin of the American Museum of Natural History 129:307–428.

Short, L. L. 1972. Hybridization, taxonomy and avian evolution. Annals of the Missouri Botanical Garden 59:447–453.

Sibley, C. G., and L. L. Short, Jr. 1959. Hybridization in the buntings (*Passerina*) of the Great Plains. The Auk 76:443–463.

Sibley, C. G., and L. L. Short, Jr. 1964. Hybridization in the orioles of the Great Plains. The Condor 66:130–150.

Sibley, C. G., and D. A. West. 1959. Hybridization in the Rufous-Sided Towhees of the Great Plains. The Auk 76:326–338.

Smith, K. L., J. M. Hale, M. R. Kearney, J. J. Austin, and J. Melville. 2013. Molecular patterns of introgression in a classic hybrid zone between the Australian tree frogs, *Litoria ewingii* and *L. paraewingi*: Evidence of a tension zone. Molecular Ecology 22:1869–1883.

Soares, C., and J. C. Brito. 2007. Environmental correlates for species richness among amphibians and reptiles in a climate transition area. Biodiversity and Conservation 16:1087–1102.

Soberón, J. 1999. Linking biodiversity information sources. Trends in Ecology and Evolution 14:291.

Spehn, E. M., A. Hector, J. Joshi, M. Scherer-Lorenzen, B. Schmid, E. Bazeley-White, C. Beierkuhnlein, M. C. Caldeira, M. Diemer, P. G. Dimitrakopoulos, J. A. Finn, H. Freitas, P. S. Giller, J. Good, R. Harris, P. Högberg, K. Huss-Danell, A. Jumpponen, J. Koricheva, P. W. Leadley, M. Loreau, A. Minns, C. P. H. Mulder, G. O'Donovan, S. J. Otway, C. Palmborg, J. S. Pereira, A. B. Pfisterer, A. Prinz, D. J. Read, E.-D. Schulze, A.-S. D. Siamantziouras, A. C. Terry, A. Y. Troumbis, F. I. Woodward, S. Yachi, and J. H. Lawton. 2005. Ecosystem effects of biodiversity manipulations in European grasslands. Ecological Monographs 75:37–63.

Stockwell, D., and D. Peters. 1999. The GARP modelling system: Problems and solutions to automated spatial prediction. International Journal of Geographical Information Science 13:143–158.

Stockwell, D. R. B., and A. T. Peterson. 2002. Controlling bias in biodiversity data. Pages 537–546 *in* J. M. Scott, P. J. Heglund, M. L. Morrison, J. B. Haufler, M. G. Raphael, W. A. Wall, and F. B. Samson, editors. Predicting Species Occurrences: Issues of Accuracy and Scale. Island Press, Washington, D.C., USA

Swenson, N. G. 2006. GIS-based niche models reveal unifying climatic mechanisms that maintain the location of avian hybrid zones in a North American suture zone. Journal of Evolutionary Biology 19:717–725.

Swenson, N. G. 2008. The past and future influence of geographic information systems on hybrid zone, phylogeographic and speciation research. Journal of Evolutionary Biology 21:421–434.

Swenson, N. G., and D. J. Howard. 2005. Clustering of contact zones, hybrid zones, and phylogeographic breaks in North America. The American Naturalist 166:581–591.

Tamura, K., G. Stecher, D. Peterson, A. Filipski, and S. Kumar. 2013. MEGA6: Molecular evolutionary genetics analysis version 6.0. Molecular Biology and Evolution 30:2725–2729.

Tavares, E., and A. Baker. 2008. Single mitochondrial gene barcodes reliably identify sister-species in diverse clades of birds. BMC Evolutionary Biology 8:81.

Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Research 22:4673–4680.

Thompson, M. C., A. Ely, B. Gress, C. Otte, S. T. Patti, D. Seibel, and E. A. Young. 2011. Birds of Kansas. First Edition. University Press of Kansas, Lawrence, KS, USA.

Tognelli, M. F., S. A. Roig-Junent, A. E. Marvaldi, G. E. Flores, J. M. Lobo. 2009. An evaluation of methods for modelling distribution of Patagonian insects. Revista Chilena de Historia Natural 82:347–360.

Tsoar, A., O. Allouche, O. Steinitz, D. Rotem, and R. Kadmon. 2007. A comparative evaluation of presence-only methods for modelling species distribution. Diversity and Distributions 13:397–405.

Vacin, V. J. 1969. Black-chinned Hummingbird at Silver Lake - near Oklahoma City. Bulletin of the Oklahoma Ornithological Society 2:14–15.

Vallender, R., V. L. Friesen, and R. J. Robertson. 2007. Paternity and performance of Golden-winged Warblers (*Vermivora chrysoptera*) and Golden-winged X Blue-winged Warbler (*V. pinus*) hybrids at the leading edge of a hybrid zone. Behavioral Ecology and Sociobiology 61:1797–1807.

Van Den Bussche, R. A., R. K. Chesser, M. J. Hamilton, R. D. Bradley, C. A. Porter, and R. J. Baker. 1993. Maintenance of a narrow hybrid zone in *Peromyscus leucopus*: A test of alternative models. Journal of Mammalogy 74:832–845.

VanDerWal, J., L. Falconi, S. Januchowski, L. Shoo, and C. Storlie. 2014. SDMTools: Species distribution modelling tools: Tools for processing data associated with species distribution modelling exercises. Available from http://cran.r-project.org/web/packages/SDMTools/SDMTools.pdf.

Ward, D. F. 2007. Modelling the potential geographic distribution of invasive ant species in New Zealand. Biological Invasions 9:723–735.

Warren, D. L., R. E. Glor, and M. Turelli. 2008. Environmental niche equivalency versus conservatism: Quantitative approaches to niche evolution. Evolution 62:2868–2883.

Warren, D. L., R. E. Glor, and M. Turelli. 2010. ENMTools: a toolbox for comparative studies of environmental niche models. Ecography 33:607–611.

Warren, D. L., and S. N. Seifert. 2011. Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. Ecological Applications 21:335–342.

Weinberg, W. 1908. Über vererbungsgesetze beim menschen. Molecular and General Genetics 1:440–460.

Wesner, J. S. 2012. Predator diversity effects cascade across an ecosystem boundary. Oikos 121:53–60.

West, D. A. 1962. Hybridization in grosbeaks (*Pheucticus*) of the Great Plains. The Auk 79:399–424.

White, P. S., and L. D. Densmore. 1992. Mitochondrial DNA isolation. Pages 29–58 *in* A. R. Hoelzel, editor. Molecular Genetic Analysis of Populations: A Practical Approach. Oxford University Press, London, UK.

Yoder, J. B., and S. L. Nuismer. 2010. When does coevolution promote diversification? American Naturalist 176:802–817.

York, E. M., C. J. Butler, and W. D. Lord. 2014. Global decline in suitable habitat for *Angiostrongylus* (= *Parastrongylus*) *cantonensis*: The Role of Climate Change. PLoS ONE 9:e103831.

Zaniewski, A. E., A. Lehmann, and J. M. Overton. 2002. Predicting species spatial distributions using presence-only data: A case study of native New Zealand ferns. Ecological Modelling 157:261–280.

Zink, R. M. 1996. Species concepts, speciation, and sexual selection. Journal of Avian Biology 27:1–6.