Summer 2023

Quantitative Song Variety in Relation to Genotype in a Hybridizing Chickadee Population

Shelby Madison Palmer
Missouri State University, Palmer904@live.missouristate.edu

QUANTITATIVE SONG VARIETY IN RELATION TO GENOTYPE IN A

HYBRIDIZING CHICKADEE POPULATION

A Master's Thesis

Presented to

The Graduate College of

Missouri State University

In Partial Fulfillment

Of the Requirements for the Degree

Master of Science, Biology

By

Shelby Madison Palmer

August 2023

QUANTITATIVE SONG VARIETY IN RELATION TO GENOTYPE IN A

HYBRIDIZING CHICKADEE POPULATION

Biology

Missouri State University, August 2023

Master of Science

Shelby Madison Palmer

ABSTRACT

The Black-capped chickadee (*Poecile atricapillus*) and Carolina chickadee (*P. carolinensis*) are North American songbird species that hybridize in a narrow contact zone stretching latitudinally from New Jersey to Kansas, USA. The association between genetic ancestry and song type in this hybrid zone has been studied independently several times and found to be minimal or absent, likely due to the influence of cultural transmission on learned song in the oscine passerine clade to which the chickadees belong. Despite this, the song of both species remains remarkably distinct in allopatry, suggesting a genetic constraint on certain qualities of their broadly learned song. I conducted genetic and acoustic sampling in a small population of chickadees in the hybrid zone in western Missouri to address the question of whether song is related to genotype from a different angle than has been taken previously. I first genotyped 55 chickadees from hybrid zone, Black-capped, and Carolina populations in Missouri and Kansas to assess the local applicability of a commonly-used genotyping method for these species, and to generate genotype scores for Missouri hybrid zone chickadees. Using active recording methods, I then obtained high-volume, high-quality recordings of songs of 10 genotyped chickadees from one hybrid zone population. I used these data to generate multivariate measurements of song variety across three different dimensions for each individual. I tested how well, and in what direction, genetic ancestry predicted song variety for each of these dimensions, predicting that song variety would increase with increasing Carolina chickadee ancestry. Linear models predicting song variety in 2 and 3 dimensions from genetic ancestry had poor fit to the data, but slope values in the predicted direction. The linear model predicting song variety in 1 dimension, similar to measurements used to characterize song phenotypes in past studies, had the worst fit to the data and a slope value near 0. These results, while not conclusive enough to confidently suggest a role of genetic ancestry in song variety, provide support for the continued use of these novel multidimensional song variety measurements and offer future directions for tackling the question of the ancestry-song relationship in the chickadee hybrid zone.

KEYWORDS: birds, hybridization, genetics, molecular markers, birdsong, bioacoustics, cultural evolution

QUANTITATIVE SONG VARIETY IN RELATION TO GENOTYPE IN A

HYBRIDIZING CHICKADEE POPULATION

By

Shelby Madison Palmer

A Master's Thesis
Submitted to the Graduate College
Of Missouri State University
In Partial Fulfillment of the Requirements
For the Degree of Master of Science, Biology

August 2023

Approved:

Jay McEntee, Ph.D., Thesis Committee Chair

Laszlo Kovacs, Ph.D., Committee Member

Mark Robbins, M.Sc., Committee Member

Julie Masterson, Ph.D., Dean of the Graduate College

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

OVERVIEW

Hybridization occurs when individuals of different species mate and produce offspring of mixed ancestry (Barton & Hewitt, 1989; Endler, 1978). The outcomes of hybridization depend on several factors, including the frequency of interspecific mating and the degree of genetic differentiation of the participating species. Hybridization can occur upon secondary contact after a period of geographic isolation during which the separated populations accumulated genetic differences, thereby beginning the process of speciation. If genetic divergence is insufficient to reduce fitness in hybrids, secondary contact will result in unrestricted gene flow and eliminate genetic structure between the previously isolated populations. If sufficient genetic divergence to reduce fitness in hybrids—but not to the point of total reproductive incompatibility—has occurred, hybrids will be selected against in areas of secondary contact, and introgression across the population boundary will be limited. This selection is counterbalanced as individuals of both species disperse into areas of secondary contact and mate heterospecifically, forming hybrid zones that can be stable over long periods of time (Endler, 1978; Bronson et al., 2003).

Avian hybridization is relatively common, having been documented in 16.4% of extant wild bird species (Ottenburghs et al., 2015). Because of this, and because hybrid zones provide unique insights into temporally-expansive evolutionary processes, avian hybrid zones are frequent hosts of studies on the genetic, phenotypic, and behavioral manifestations of hybridization in birds (Bennett et al., 2021; DeRaad et al., 2023; Semenov et al., 2021; Toews et al., 2016; Wheatcroft & Qvarnström, 2017). One such hybrid zone is that of the Black-capped chickadee (*Poecile atricapillus*) (hereafter BCCH) and Carolina chickadee (*Poecile carolinensis*) (hereafter CACH), congeneric members of the songbird family Paridae which hybridize in a

1

narrow latitudinal zone of range overlap bisecting the eastern United States from New Jersey to Kansas (Robbins et al., 1986; Bronson et al., 2005). BCCH and CACH are thought to have diverged 2 million years ago during a period of geographic isolation (Gill et al., 1993) and now hybridize frequently in a relatively stable zone of secondary contact. Recent genomic analyses of the BCCH/CACH hybrid zone cline center suggest that the zone is moving northward, with the eastern portion experiencing more rapid movement and widening (Taylor et al., 2014b; Bronson et al., 2005) than the western portion (Alexander et al., 2022). Climate change is thought to play a part in this movement, as increasing temperatures expand the suitable range for the less cold-adapted CACH (Taylor et al., 2014a; Taylor et al., 2015; Alexander et al., 2022). The social dominance of CACH over BCCH could act in tandem with climate change to drive the northward movement of the hybrid zone (Reudink et al., 2006).

The latitudinal expansiveness of the BCCH/CACH hybrid zone, its northward movement, and the abundance and bold nature of both species make this study system ideal for investigating a broad range of ecological and evolutionary questions. Indeed, the BCCH/CACH hybrid zone has played host to many such studies, including some exploring the dynamics of song, an acoustically-complex vocalization, in genetically-admixed populations of chickadees. Development of song in the oscine passerine clade of songbirds—including the chickadees—has a strong learned component, making it less constrained by an innate, inherited neural template and more sensitive to the forces of cultural transmission than that of taxa that do not exhibit learning (Beecher & Brenowitz, 2005; Marler, 1990; Soha & Marler, 2000). Hybrid zones in which both genetic and vocal admixture are present provide opportunities to elucidate the relative roles of learning and genetics in song development. The majority of studies focusing on song in the BCCH/CACH hybrid zone have been implemented in its eastern portion (Curry et al.,

2

2007; Sattler et al., 2007; Wright Nelson, 2016; Abbrescia, 2021), although Braun & Robbins (1986) and Robbins et al. (1986) conducted thorough genetic and acoustic sampling in west-central Missouri. Regardless of location, a lack of association between genotype and song has been the primary finding of all of these studies, supporting the hypothesis that song in this taxon is freely learned across the species boundary and is relatively unconstrained by genetics. This has been the consensus of studies on song in other oscine passerine hybrid zones as well (Emlen et al., 1975; Kenyon et al., 2011; Kenyon et al., 2017). However, the distinctiveness of BCCH and CACH song in allopatry and some limitations of methodological approaches to acoustic analysis in past studies raise the possibility that yet-uninvestigated characteristics of BCCH/CACH song are under genetic constraint. The identification of such characteristics would enhance our understanding of how song develops in oscine passerines, and could change the way we approach quantifying variation in birdsong.

This study aims to investigate the relationship between ancestry and quantitative measures of individual song variety in a small BCCH/CACH hybrid zone population in western Missouri. In Chapter 1, I implement a molecular marker protocol commonly used to genotype BCCH, CACH, and their hybrids in the eastern portion of the hybrid zone (McQuillan et al., 2017), assess the applicability of this method in the western portion of the zone, and generate genotype scores for Missouri hybrid zone chickadees. In Chapter 2, I apply a novel method for objectively quantifying individual-level acoustic variation in this BCCH/CACH hybrid zone population, and test how well genetic ancestry predicts different measurements of song variation.

CHAPTER 1: TESTING SPECIES-DIAGNOSTIC MOLECULAR MARKERS AND CHARACTERIZING THE GENETIC BACKGROUND OF MIXED-SINGING CHICKADEES

Introduction

The selection of molecular methods with which to conduct a hybrid zone study depends largely upon the study's scope and aims. Increasingly affordable whole-genome sequencing and the advent of next-generation genomic techniques enable high-precision estimates of hybrid zone dynamics using cline analysis (Taylor, 2014a; Wagner, 2020; Alexander et al., 2022), as well as identification of highly-divergent regions of the genome (Toews et al., 2016; Brelsford et al., 2017; Taylor, 2014b) and detection of the presence and direction of introgression (Alexander et al., 2022). Other studies have used molecular markers to address questions about the biogeography of the parental species (Gill et al., 1993) and test for association between ancestry and certain phenotypic or behavioral traits (McQuillan et al., 2018; Van Huynh & Rice, 2019; Abbrescia, 2021; Sattler et al., 2007).

Studies using molecular markers to investigate hybrid zones must take into account the specific goals of the study when selecting markers, since various characteristics of the chosen subset of loci could have a dramatic effect on downstream results. Allele frequency divergence is an important factor: a set of highly-divergent species-diagnostic molecular markers may be effective for diagnosing admixed and parental individuals in a hybrid zone, but would likely result in underestimation of levels of introgression (Yuri et al., 2009). The procedural details behind the establishment of the markers are also worth considering; for example, molecular marker methods meant to be species-diagnostic are sometimes established using small,

geographically-invariant subsets of individuals which may not be genetically representative of the entire species. Linkage disequilibrium (hereafter LD), the nonrandom association between multiple loci, can also affect the information ascertained from a set of molecular markers. The increased likelihood of certain loci to be inherited together through generations due to LD ascribed to physical linkage (sharing a chromosome) introduces a degree of non-independence among loci employed as markers. For markers that are physically linked, LD increases with increasing marker proximity and decreases with recombination rate (Slatkin, 2008). Additionally, in hybrid zones, LD in both physically linked and unlinked markers is heightened due to substantial allele frequency differences between the two hybridizing species, a phenomenon known as admixture linkage disequilibrium (hereafter admixture LD) (Stephens et al., 1994; Falush et al., 2003). The factors mentioned above, among others, highlight the importance of becoming familiar with one's study system and selected molecular marker method when investigating the genetic composition of hybrid zones. This familiarity allows us to account for potentially confounding factors in our methods and analyses, preventing us from arriving at inaccurate results and misleading conclusions.

A set of 10 species-diagnostic autosomal SNP markers for BCCH and CACH were obtained by McQuillan et al. (2017) based on transcriptome databases from 5 BCCH from Ithaca, New York and 5 CACH from Baton Rouge, Louisiana. These markers were used to develop a quick, inexpensive PCR-RFLP genotyping method for species diagnosis of BCCH and CACH (McQuillan et al., 2017). This method has since been employed to diagnose parental BCCH and CACH and to generate genotype scores for hybrids in several recent studies investigating aspects of the BCCH/CACH hybrid zone in eastern Pennsylvania (McQuillan et al., 2018; Van Huynh & Rice, 2019; Rice et al., 2021; Abbrescia, 2021; Driver et al., 2022).

Every step of this method is easily executed in a basic molecular lab, making it desirable for short-term studies requiring multiple rounds of genetic data collection. However, the markers it employs were derived from BCCH and CACH representing a single population per species and were only verified to be species-diagnostic for individuals also from New York and Louisiana (McQuillan et al., 2017). Taking into account both the convenience and potential drawbacks of the method established by McQuillan et al. (2017), in addition to using these markers to obtain genotype scores for putative hybrid chickadees in Missouri, I applied the method to parental BCCH and CACH from Missouri and Kansas to determine their effectiveness in species diagnosis in Midwestern populations of BCCH and CACH. All but one of the 10 markers from McQuillan et al. 2017 are physically linked to at least one other; combined with the presence of admixture LD, this allows for the possibility of levels of LD resulting in non-independent data from markers. To account for these possibilities, I calculated differences in LD coefficients for physically linked and unlinked pairs of markers based on the allele frequencies found in my putative hybrid zone populations to assess their power to estimate ancestry fractions.

Methods

Capture Site Selection. From May – August of 2021, I located potential chickadee capture sites by driving repeatedly across the BCCH/CACH range boundary (location approximated from Robbins et al., 1986 and Robbins, pers. comm.) in west-central Missouri and searching for chickadees who displayed mixed singing and/or species-intermediate song. For legal reasons, the search was narrowed to tracts of public land. Upon finding a public use area with appropriate habitat, I looked and listened for chickadees, attempting to prompt singing using playback of BCCH/CACH vocalizations from the Merlin app (Cornell Lab of Ornithology)

6

broadcast from a JBL Flip5 speaker paired to a smartphone. When singing chickadees appeared, I made preliminary song recordings with a Sound Devices MixPre 3 II audio recorder/Sennheiser ME 62 microphone mounted in a parabolic dish for 5-10 minutes or until the individual left the area. If the individual sang atypically of either species, I attempted to follow it for 30-45 minutes while taking location data to facilitate re-location and subsequent capture. After this preliminary search, I chose to focus my sampling efforts in two public parks in Henry County, Missouri: Sparrowfoot Park and Clinton City Park. I chose these locations for ease of access, abundance of chickadees, and presence of mixed/unusual song types.

Blood Sampling and Color Banding. I captured chickadees at the pre-determined sites during September-October 2021, March-August 2022, and February-March 2023 using mist-nets and playback of song and call vocalizations of both BCCH and CACH. Individuals were fitted with a unique 3-color combination of plastic bands and one USGS-issued aluminum band. A 10-40 uL blood sample was collected in a capillary tube via brachial venipuncture and immediately deposited in 1 mL Queen's lysis buffer (Seutin et al., 1991). If the bird seemed in good condition, morphological measurements were taken following color banding and blood sampling. After processing, birds were immediately released and monitored for 2-5 minutes to assess recovery from handling. Blood samples in Queen's lysis buffer were stored at 5ºC prior to DNA extraction. This protocol was approved by the Institutional Animal Care and Use Committee on January 5, 2021 and received Approval #2021-22 (Appendix A).

PCR-RFLP Genotype Scoring. I extracted DNA from each blood sample using a Qiagen DNeasy Blood & Tissue kit and checked the DNA concentration of each extraction on an Implen P-330 Nanophotometer. Extractions that consistently returned a concentration reading <10 ng/µL were discarded, and the DNA extraction process was attempted again for those

individuals. If DNA concentration was <10 ng/µL for the second extraction attempt, that

individual was excluded from genotyping.

To obtain individual genotypes, I used the following PCR-RFLP method for 9 of the 10

putatively species-diagnostic loci in BCCH/CACH described by McQuillan et al. 2017

(Appendices B & C). The locus named c0p171 in McQuillan et al. 2017 consistently failed to

amplify and was excluded from analyses. PCRs were run in 10 µL volumes with the following

contents: 0.2 µL F primer (Appendix B), 0.2 µL R primer (Appendix B), 2.0 µL PCR-grade

water, 5.0 µL 2X DreamTaq master mix (Thermo Fisher Scientific), and 2.6 µL DNA template.

For reactions that initially failed to amplify, the PCR-grade water was replaced with an equal

volume of 3.0% dimethyl sulfoxide (DMSO) in subsequent runs. PCR was run in a T100 BioRad

thermal cycler under the following conditions: 3:00 at 95ºC, 35 cycles of 0:30 at 95ºC, 0:30 at T_a

(Appendix B), 1:00 at 72ºC, and 7:00 at 72ºC. I confirmed the presence of each PCR product

using agarose gel electrophoresis (2.0% agarose gels run at 100V for 30 minutes) alongside a

100-bp DNA ladder. For successfully-amplified samples that produced a bright band, the

remaining product was combined with 10 units of a restriction enzyme designated for the target

locus (Appendix C) and incubated according to the protocol provided by the manufacturer.

Digestion products were electrophoresed on 2.0% agarose gels at 100V for 45 minutes alongside

a 100-bp DNA ladder and visualized under UV light (Figure 1.1). DNA fragment sizes were

compared with species-specific base pair lengths for the digestion product of each locus

(Appendix C). In accordance with the observed fragment sizes, each individual received a score

of either BC/BC, BC/CA, or CA/CA for each of the 9 loci. For use in downstream analyses,

these scores were input as individual alleles in a .csv file in the format required by the program

STRUCTURE (Pritchard et al., 2000) with a BCCH allele coded as 1 and a CACH allele coded as 2.

To determine the applicability of the McQuillan et al. (2017) PCR-RFLP method for genotyping individuals from the hybrid zone in Missouri, I performed the methods detailed above using DNA from 11 BCCH and 10 CACH from eastern Kansas, west-central Missouri, and southern Missouri (Figure 1.2). Blood samples of one BCCH and one CACH captured ~50 km. north and ~150 km. south, respectively, of the hybrid zone were included in the protocol detailed above. Additionally, liver tissue of 10 BCCH and 10 CACH individuals were loaned from the University of Kansas Museum of Natural History (Appendix D). DNA for each individual was extracted from <0.1 g of excised and chopped tissue using a Qiagen DNeasy Blood & Tissue Kit. One CACH sample was excluded from analyses due to an insufficient DNA concentration in the DNA extraction. PCR-RFLP was then performed and scored identically to the method described above.

STRUCTURE Analysis. I used the Bayesian clustering program STRUCTURE (Pritchard et al., 2000) to generate genotype scores for each sampled individual (n=34 HZ, n=11 BCCH, n=10 CACH). When sequence data of sampled individuals are input, STRUCTURE uses Markov chain-Monte Carlo (MCMC) estimation to assign individuals to K populations, resulting in a score Q of probability of assignment to each of K populations for each individual. I used the Q scores corresponding to the probability of assignment to the CACH-like population cluster for song variety analyses in Chapter 2. Prior to running SRUCTURE, I used the Basic Local Alignment Search Tool on the National Center of Biotechnology Information website (https://blast.ncbi.nlm.nih.gov/Blast.cgi) to align the sequences for each of the markers from McQuillan et al. (2017) to the zebra finch reference genome (following the methods from
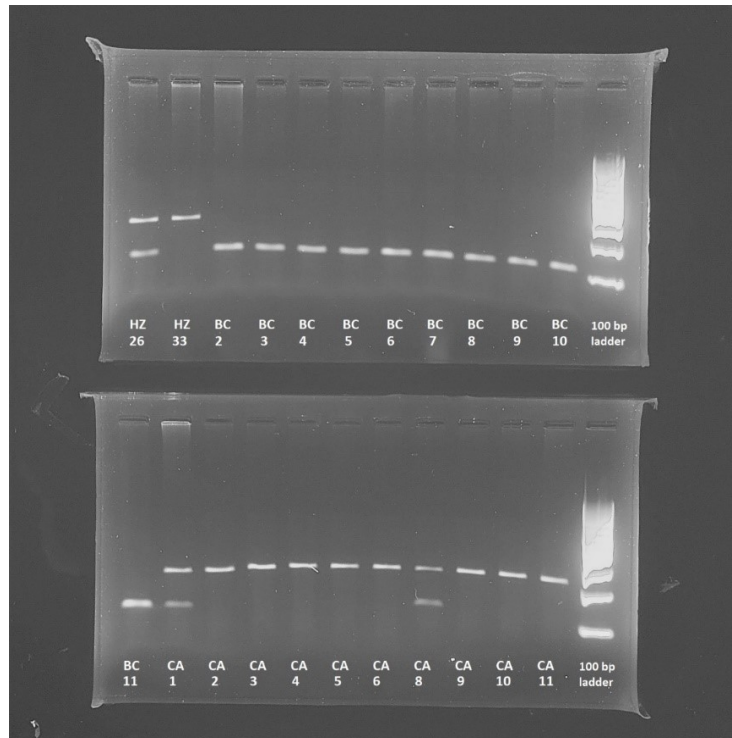
Figure 1.1. Agarose gel electrophoresis of the PCR-RFLP product containing the marker c0p628. The restriction enzyme BsaAI cuts the 217- base pair PCR product into 107 and 110 base pair fragments for the BCCH allele, and leaves the product uncut for the CACH allele. For this marker, heterozygous individuals produce both cut and uncut DNA fragments, which appear as two distinct bands (HZ26, CA1, and CA8 above).
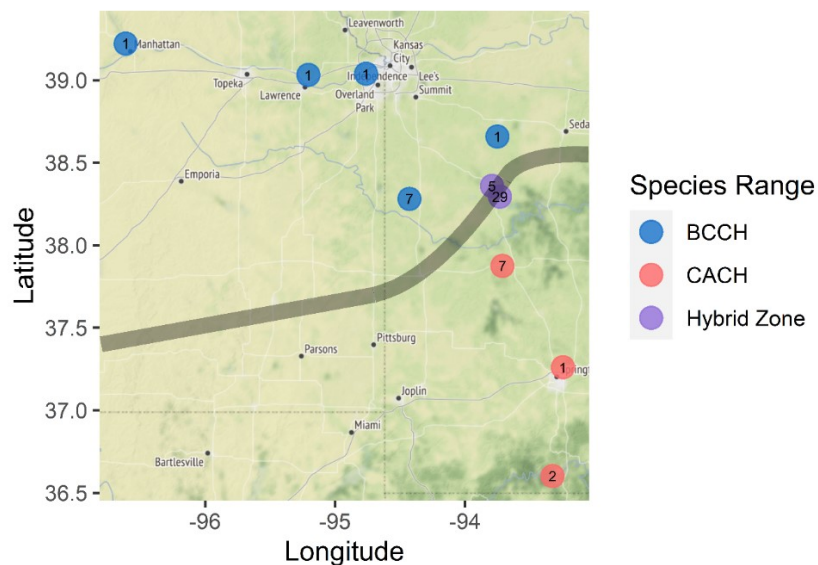
Figure 1.2. A map displaying the approximate location of the BCCH/CACH hybrid zone (gray line) through eastern Kansas and western Missouri and sampling localities of parental and putative hybrid individuals. The numbers overlaid on the points are counts of individuals sampled at that location.

McQuillan et al., 2017) and recorded both the chromosome number and the position on the chromosome (in number of nucleotides) for each. This allowed me to obtain physical inter-marker distances for consecutive markers on the same chromosome. The STRUCTURE input file contained the following information: individual ID codes, PCR-RLFP scores for each individual, sampling location information for each individual, marker ID codes, and physical distance (number of nucleotides) between consecutive markers that mapped onto the same chromosome. I ran STRUCTURE with the following parameters: $K=2$, burnin runs = 10,000, runs = 100,000. STRUCTURE's admixture ancestry model, the model typically used to assess population structure when admixture is known to occur, assumes Hardy-Weinberg equilibrium and linkage equilibrium. To account for the possibility of LD, I selected the linkage ancestry model (Falush et al., 2003), which includes provided measures of genetic or physical distance between loci as a model parameter.

Allele Frequencies by Locality. Using the genotype scores from the PCR-RFLP procedure, I calculated the frequencies of the BCCH-designated and CACH-designated alleles for each marker from McQuillan et al. 2017 for each sampling locality (Table 1.1, Figure 1.3).

Testing for Linkage Disequilibrium. Using the allele frequencies for hybrid zone individuals only, I calculated the LD coefficients D, D', and r^2 among all possible pairs (n = 36) of the 9 markers using the R package genetics (v.4.3.1, R Core Team, 2023; Warnes et al., 2021). D is calculated between two loci with the equation $D = p_{AB} - p_A p_B$, where p_{AB} is the population-level frequency of alleles A and B co-occurring at two loci in a gamete (i.e., the frequency of the AB haplotype) and $p_A p_B$ is the product of the population-level frequencies of alleles A and B at two loci. Because haplotype frequencies cannot be obtained from diploid genotypes without gametic phase data, the genetics package estimates p_{AB} using maximum likelihood (Warnes et

11

Table 1.1. Frequencies of the species-diagnostic BCCH and CACH alleles designated in McQuillan et al. 2017 for each sampling locality.

BCCH allele frequency	CACH allele frequency	# of individuals sampled	Species range	City/Area	County	State
1	0	1	BCCH	Manhattan	Riley	KS
1	0	1	BCCH	Lawrence	Douglas	KS
1	0	1	BCCH	Lake Quivira	Johnson	KS
0.944	0.0556	1	BCCH	Leeton	Johnson	MO
0.952	0.0476	7	BCCH	Butler Lake	Bates	MO
0.756	0.244	5	Hybrid Zone	Clinton	Henry	MO
0.216	0.784	29	Hybrid Zone	Sparrowfoot Park	Henry	MO
0.0714	0.929	7	CACH	Bird Song CA	St. Clair	MO
0.111	0.889	1	CACH	Springfield	Greene	MO
0.0278	0.972	2	CACH	Table Rock Lake	Stone	MO
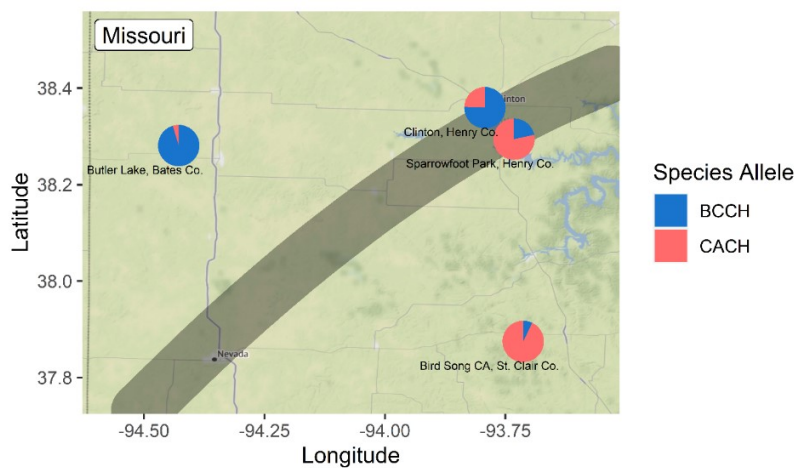
Figure 1.3. A map displaying the frequencies of alleles designated as BCCH- and CACH-specific by McQuillan et al. (2017) across the 9 genotyped markers for each sampling locality represented by 5 or more individuals. The gray band marks the approximate location of the hybrid zone.

al., 2021). D' is the observed D divided by the maximum possible value of D given the frequencies of alleles A and B at the two loci. r^2 is a correlation coefficient of alleles A and B at the two loci and is derived by the following equation: $r^2 = D^2/p_A(1 - p_A)p_B(1 - p_B)$. I ran two-sample Welch's t-tests to determine whether LD coefficients differed significantly between groups of physically linked and physically unlinked markers.

Results

    Genetic Composition of Hybrid Zone Populations. Genetic admixture was present at both of the putative hybrid zone sampling localities (Table 1.1, Figs. 1.3 & 1.4). Chickadees captured at Sparrowfoot Park, the primary hybrid zone site, were predominantly a mix of CACH and CACH-like advanced generation hybrids, along with a few likely first-generation hybrids and one BCCH (mean probability of assignment to CACH = 0.80 ±0.24 sd, frequency of CACH allele = 0.78, n = 29). Chickadees captured at Clinton City Park, the secondary hybrid zone site, were a combination of BCCH, BCCH-like advanced generation hybrids, and one CACH-like advanced generation hybrid (mean probability of assignment to CACH = 0.25 ±0.27 sd, frequency of CACH allele = 0.24, n = 5).

    Genetic Composition of Non-Hybrid Zone Populations. The presence of at least 1 CACH allele was detected in 6 of 11 allopatric BCCH, producing STRUCTURE Q scores and allele frequencies suggestive of genetic admixture in local BCCH populations, although still notably less than in hybrid zone populations (mean probability of assignment to CACH = 0.023 ±0.020 sd, frequency of CACH allele = 0.035, n = 11) (Table 1.1, Figs. 1.3 & 1.4). Interestingly, CACH alleles were only detected at the two Missouri sampling localities; chickadees from Kansas appeared to be pure BCCH. The presence of at least 1 CACH allele was detected in 7 of

10 CACH, producing STRUCTURE Q scores and allele frequencies suggestive of genetic admixture in local CACH populations, although also less than in hybrid zone populations (mean probability of assignment to CACH = 0.96 ±0.056 sd, frequency of CACH allele = 0.93, n = 10) (Table 1.1, Figs. 1.3 & 1.4). One of the markers, c0p283, was responsible for a disproportionate amount of the heterospecific allele presence in populations of both parental species, accounting for 47% of the heterospecificity observed across the 9 markers outside the hybrid zone.
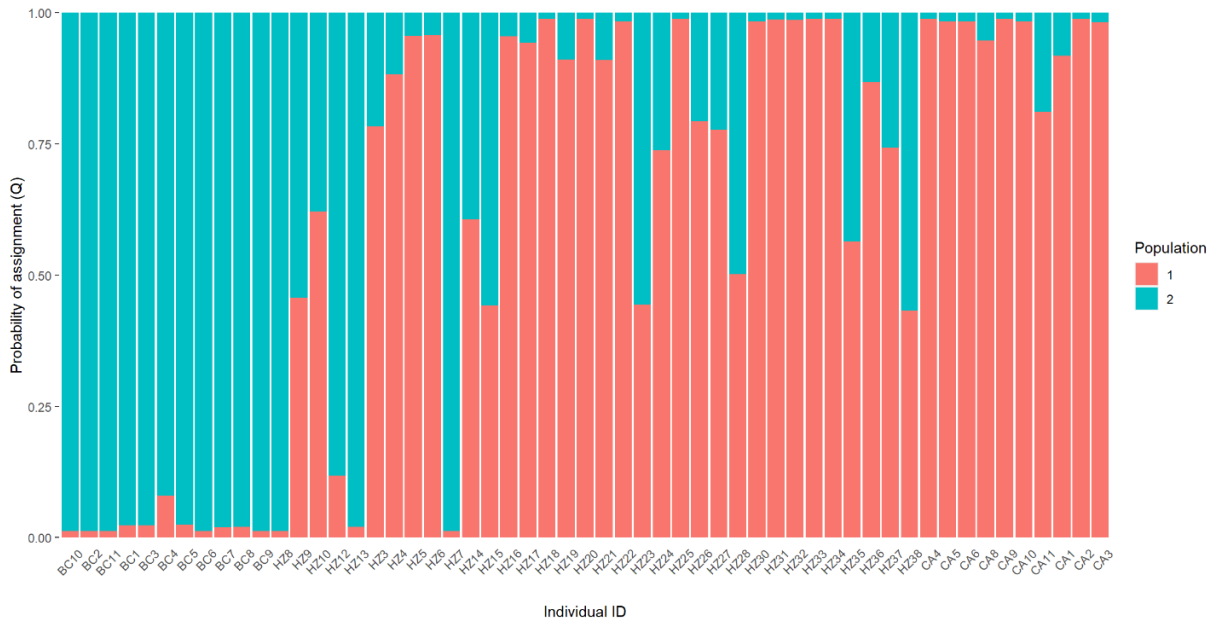
Figure 1.4. Probabilities of assignment Q to each of K=2 populations of 11 BCCH (prefix BC), 10 CACH (prefix CA), and 34 chickadees of unknown ancestry (prefix HZ) generated by STRUCTURE. Individuals are arranged by sampling locality, starting from the left with BCCH sampled farthest away from the hybrid zone, and ending on the right with CACH sampled farthest away from the hybrid zone (distances of BCCH and CACH sampling localities from the hybrid zone estimated from Alexander et al. 2022).

Linkage Disequilibrium in McQuillan et al. (2017) Markers. There was not a significant difference between physically linked and physically unlinked marker pairs for any of the LD coefficients (D: t(6.52) = -1.64, p = 0.149; D': t(6.86) = -1.42, p = 0.199; r^2: t(6.2) = 0.54, p = 0.605) (Figure 1.5). One physically linked marker pair, c0p356 and c0p628, had

consistently high values across all coefficients (Figure 1.6).
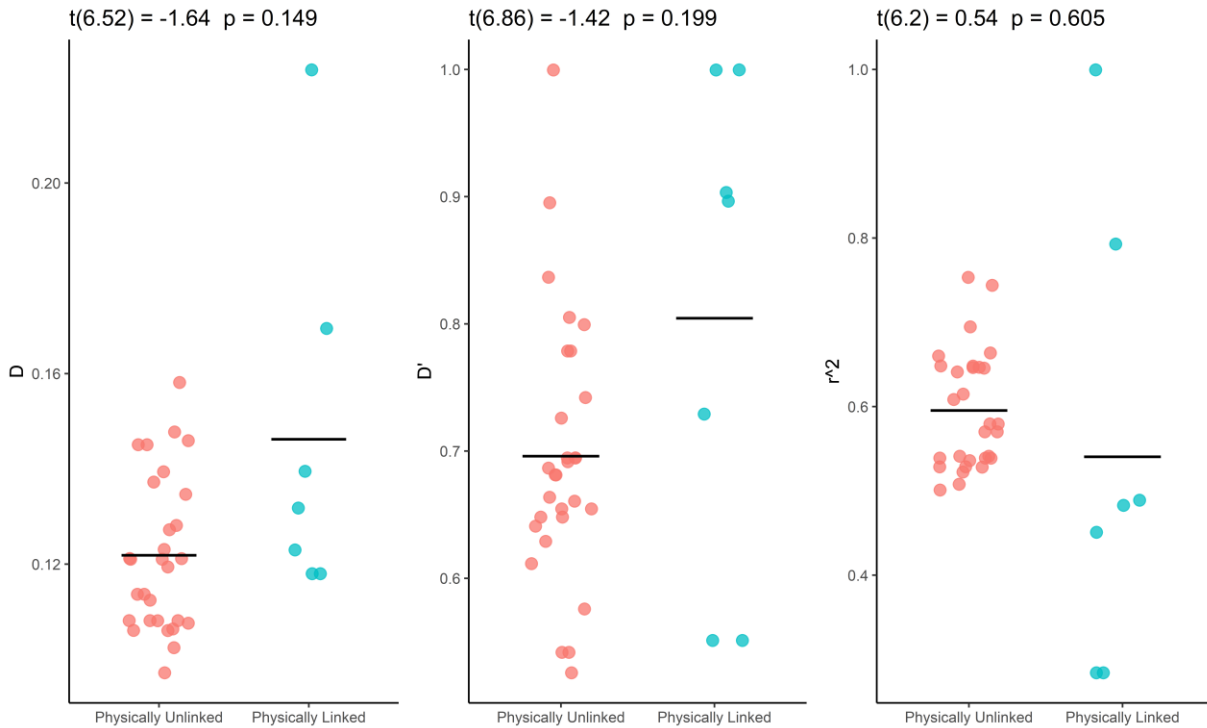
Figure 1.5. Strip charts of pairwise measures of the linkage disequilibrium coefficients *D*, *D'*, and r^2 between each physically-linked and physically-unlinked pair of markers. Text above each plot gives results of Welch's t-tests for physically linked and unlinked marker groups. Group means are marked by black horizontal lines.

Discussion

The genetic admixture detected at my two putative hybrid zone sites is expected given the approximate location of the hybrid zone in western Missouri as reported by Braun & Robbins (1986) and Robbins et al. (1986) and accounting for the displacement of the cline center of this part of the hybrid zone ~5 km to the northwest (Alexander et al., 2022). Although hybridization was occurring at both sites, neither appeared to be at the genetic cline center of the hybrid zone; the genetic composition of Sparrowfoot Park was disproportionately CACH, and that of Clinton City Park was disproportionately BCCH, despite their mere 8-km separation. This degree of
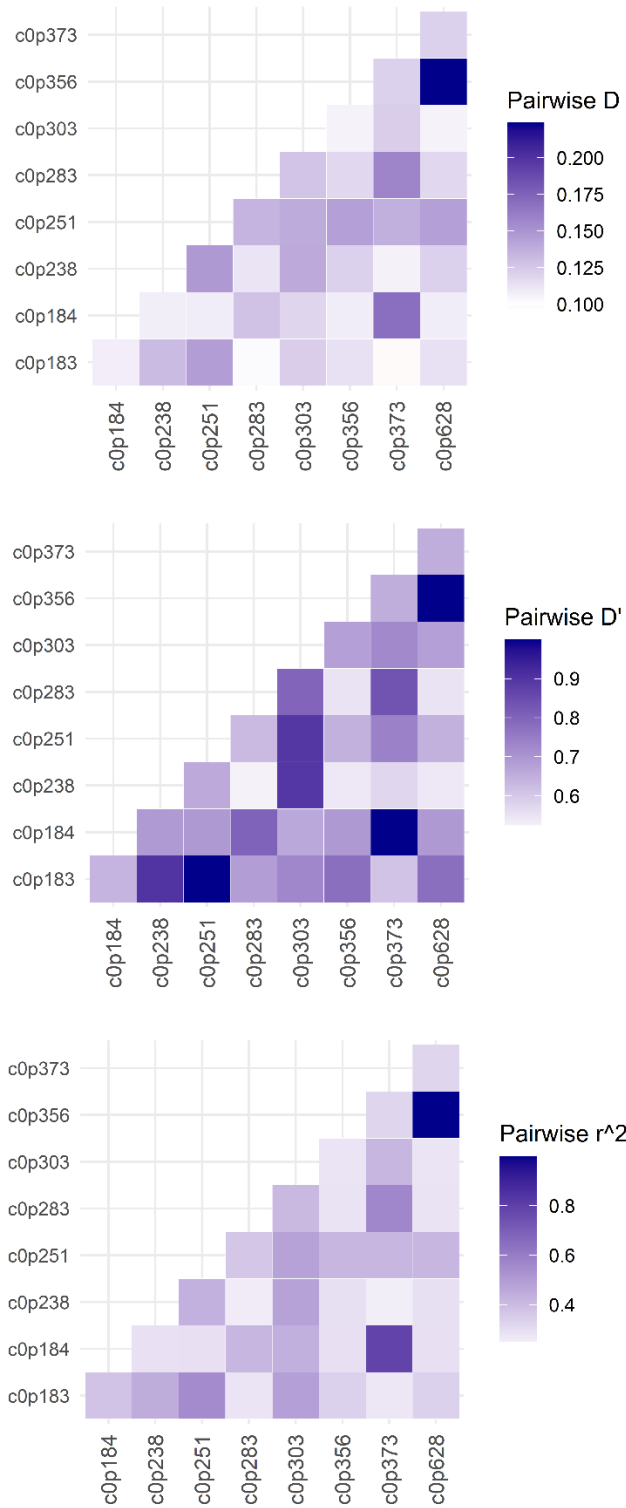
Figure 1.6. Heat maps displaying magnitudes of the LD coefficients *D*, *D'*, and r^2 between each pair of markers.

genetic difference between two spatially proximate sites is expected due to the narrow (15 km) width of the hybrid zone in western Missouri (Robbins et al., 1986; Alexander et al., 2022).

The presence of heterospecific alleles in Missouri BCCH and CACH allopatric populations implies that either the extent of introgression across the hybrid zone is greater than previously estimated, or the species-diagnostic markers described in McQuillan et al. 2017 are not fixed in chickadee populations in Missouri. The former implication would require either that the historical 15-km estimate of the hybrid zone width was overly conservative, or that it has expanded significantly in both directions since the 1980's. The 15-km estimate was originally based on song and morphology intermediacy (Robbins et al., 1986) and was recently re-analyzed using a RAD-Seq generated 11,669-SNP dataset including both contemporary (2016) and historical data (Alexander et al., 2022). This analysis showed that the historical zone likely extended a few kilometers farther to the northwest than estimated by Robbins et al. (1986) due to the sampling of individuals that were vocally BCCH-like, but genetically admixed. Alexander et al. (2022) were limited in investigating potential expansion of the hybrid zone width due to lack of sampling into the contemporary range of BCCH. However, they were able to demonstrate northwestward movement of the trailing edge of the hybrid zone, suggesting that the zone is moving, but not significantly expanding (Alexander et al., 2022). These results suggest that the hybrid zone width was not underestimated enough to explain the presence of heterospecific alleles well outside the zone, nor has it expanded significantly since the 1980's. Thus, the possibility that the McQuillan et al. (2017) markers are not fixed across the ranges of BCCH and CACH seems the most likely explanation for their presence in heterospecific chickadee populations in Missouri.

The lack of difference in LD coefficients between pairs of physically linked and

physically unlinked markers suggest that LD due to physical linkage did not significantly impact the independence of genotype data obtained from the McQuillan et al. (2017) markers. Due to the potential effect of small sample size on the statistical test applied, I interpret this result with some caution. Notably, the marker pair of c0p356 and c0p228, both of which mapped onto chromosome 21 of the zebra finch reference genome, have nearly the maximum value of both D and r^2, suggesting the possibility of LD between these two loci. Several marker pairs, both physically linked and unlinked, have nearly maximal values of D'. This could result from admixture LD between loci unrelated to physical linkage. Overall, with the exception of the aforementioned marker pair on chromosome 21, physical linkage does not appear to drive a disproportionate amount of the LD present among the McQuillan et al. (2017) markers.

Taken together, these results demonstrate the importance of exercising caution when using molecular markers, even in a study system for which they were specifically designed. Molecular marker methods are convenient and effective ways to obtain genetic data, but only when the aims of the study align with the intended use of the method, the latter of which is not always clearly defined. For example, the method described by McQuillan et al. (2017) was tested on chickadees from a hybrid zone population in Pennsylvania and has subsequently been used exclusively in and near the hybrid zone in eastern Pennsylvania (with the exception of this study). It is possible, but not explicitly stated, that this method was designed for use in the eastern portion of the chickadee hybrid zone specifically, since the majority of research on the BCCH/CACH hybrid zone is conducted there. Additionally, the marker from McQuillan et al. (2017) that disproportionately contributed to heterospecific allele presence outside the hybrid zone in my study, c0p283, was excluded from genotyping in two previous studies employing this method without an explanation (Rice et al., 2021; Driver et al., 2022). Perhaps the authors had

some knowledge that c0p283 was heterozygous in parental populations where it should have been fixed or absent, but this is not stated in the literature. As general best-practice rules, researchers using molecular markers for genotyping should thoroughly research past applications of prospective methods and apply the chosen method to control samples for which ancestry is known. The former prevents mis-selection of an inappropriate set of markers, while the latter allows us to catch and correct for any markers that deviate from expected allele frequencies in parental populations.

References

Abbrescia, A. (2021). Dawn chorus male song patterns in relation to ancestry in the Black-capped Chickadee × Carolina Chickadee hybrid zone. *Villanova Research Journal*, *3*, 47-57.

Alexander, A., Robbins, M. B., Holmes, J., Moyle, R. G., & Peterson, A. T. (2022). Limited movement of an avian hybrid zone in relation to regional variation in magnitude of climate change. *Molecular Ecology*, *31*(24), 6634–6648.

Brelsford, A., Toews, D. P. L., & Irwin, D. E. (2017). Admixture mapping in a hybrid zone reveals loci associated with avian feather coloration. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1866), 20171106.

Braun, M. J., & Robbins, M. B. (1986). Extensive Protein Similarity of the Hybridizing Chickadees Parus atricapillus and P. carolinensis. *The Auk*, *103*(4), 667–675.

Driver, R. J., Ferretti, V., Burton, E. S., McCoy, M. W., Cornell Duerr, K. L., Curry, R. L. (2022). Spatiotemporal Variation in Hatching Success and Nestling Sex Ratios Track Rapid Movement of a Songbird Hybrid Zone. *The American Naturalist*, 200(2), 264–274.

Falush, D., Stephens, M., & Pritchard, J. K. (2003). Inference of Population Structure Using Multilocus Genotype Data: Linked Loci and Correlated Allele Frequencies. *Genetics*, *164*(4), 1567–1587.

Gill, F. B., Mostrom, A. M., & Mack, A. L. (1993). Speciation in North American Chickadees: I. Patterns of MT DNA genetic divergence. *Evolution*, *47*(1), 195–212.

McQuillan, M. A., Huynh, A. V., Taylor, S. A., & Rice, A. M. (2017). Development of 10 novel SNP-RFLP markers for quick genotyping within the black-capped (Poecile atricapillus) and

Carolina (P. carolinensis) chickadee hybrid zone. *Conservation Genetics Resources*, *9*(2), 261–264.

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of Population Structure Using Multilocus Genotype Data. *Genetics*, *155*(2), 945–959.

R Core Team (2023). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rice, A. A., Curry, R. L., & Weckstein, J. D. (2021). Haemosporidian prevalence and community composition vary little across a chickadee hybrid zone. *Ornithology*, *138*(3), 1-13.

Robbins, M. B., Braun, M. J., & Tobey, E. A. (1986). Morphological and vocal variation across a contact zone between the chickadees Parus atricapillus and P. carolinensis. *The Auk*, *103*(4), 655–666.

Seutin, G., White, B. N., & Boag, P. T. (1991). Preservation of avian blood and tissue samples for DNA analyses. *Canadian Journal of Zoology*, *69*(1), 82–90.

Slatkin, M. (2008). Linkage disequilibrium—Understanding the evolutionary past and mapping the medical future. *Nature Reviews Genetics*, *9*(6), 477–485.

Stephens, J. C., Briscoe, D., & O'Brien, S. J. (1994). Mapping by Admixture Linkage Disequilibrium in Human Populations: Limits and Guidelines. *American Journal of Human Genetics*, *55*, 809–824.

Taylor, S. A., White, T. A., Hochachka, W. M., Ferretti, V., Curry, R. L., & Lovette, I. (2014). Climate-Mediated Movement of an Avian Hybrid Zone. *Current Biology*, *24*(6), 671–676.

Van Huynh, A., & Rice, A. M. (2019). Conspecific olfactory preferences and interspecific divergence in odor cues in a chickadee hybrid zone. *Ecology and Evolution*, *9*(17), 9671–9683.

Taylor, S. A., Curry, R. L., White, T. A., Ferretti, V., & Lovette, I. (2014). Spatiotemporally consistent genomic signatures of reproductive isolation in a moving hybrid zone: Genomics of a moving hybrid zone. *Evolution*, *68*(11), 3066–3081.

Wagner, D. N., Curry, R. L., Chen, N., Lovette, I. J., & Taylor, S. A. (2020). Genomic regions underlying metabolic and neuronal signaling pathways are temporally consistent in a moving avian hybrid zone. *Evolution*, *74*(7), 1498–1513.

Warnes, G., Gorjanc, G., Leisch, F., & Man, M. (2021). *genetics: Population Genetics*. R package version 1.3.8.1.3.

Yuri, T., Jernigan, R. W., Brumfield, R. T., Bhagabati, N. K., & Braun, M. J. (2009). The effect of marker choice on estimated levels of introgression across an avian (Pipridae: Manacus) hybrid zone. *Molecular Ecology*, *18*(23), 4888–4903.

CHAPTER 2: QUANTIFYING SONG VARIETY AND ITS RELATIONSHIP TO ANCESTRY IN A HYBRID ZONE CHICKADEE POPULATION

Introduction

Vocal communication is widespread across avian taxa. In the order Passeriformes, the songbirds, vocalizations are broadly categorized into calls and song (Catchpole & Slater, 2008). Calls are highly context-specific and tend to have phylogenetically conserved acoustic qualities. Song, produced mostly by males, is more acoustically complex and primarily functions in territory defense and mate acquisition (Nowicki & Searcy, 2014). The developmental drivers behind song vary between the two major subdivisions of the Passeriformes. In sub-oscine passerines, song develops according to a rigid neural template with a minimal role of cultural influence (Kroodsma & Konishi, 1991). Conversely, in oscine passerines, this neural template is more flexible, and learning from conspecifics is a major driver of song development (Beecher & Brenowitz, 2005; Marler, 1990; Soha & Marler, 2000). The plasticity of song development in oscine passerines has allowed for the evolution of extensive vocal variation across this taxon, as both genetics and geography play a part in song transmission (Catchpole & Slater, 2008; Marler, 1990). Also important in shaping this variation is the heavy influence of sexual selection on song evolution, resulting in the propagation of various song qualities that honestly indicate fitness in males (Catchpole & Slater, 2008; Goller, 2021), a process also thought to accelerate speciation (Mason et al., 2017). The sheer amount of variation present in birdsong and the multimodality of its development have prompted many to study how this fascinating signal evolves. Hybrid zones between closely-related oscine passerine species, especially those which have evolved

acoustically-distinct songs, provide opportunities to investigate the relative roles of cultural and genetic evolution in the development of birdsong.

BCCH and CACH sing relatively short, whistled songs that are distinct throughout most of their respective ranges (Figure. 2.1). BCCH sing a two-note song with a drop in frequency between the first and second notes, often with distinct amplitude modulation in the second note; it is often described as sounding like "fee-bee" or "fee-bee-ee." BCCH song is remarkably stereotyped across its extensive range, with dialectal variation concentrated in populations at its outer extremes (Kroodsma et al., 1999). There is some support for stereotypy in BCCH song acting as an indicator of fitness: a study investigating the relationship between social dominance and consistency in certain qualities of song in BCCH found that males who maintained a constant inter-note frequency interval across song bouts were likely to rank higher in their winter flock dominance hierarchy (Christie et al., 2004). The most common version of CACH song is a four-noted whistled "see-bee-see-bay," higher-pitched than that of the BCCH, with notes at alternating high and low frequencies. Unlike BCCH, CACH shows considerable song variation both regionally and at the individual level (Ward, 1966; Keleman et al., 2015) although studies characterizing song variation in CACH across its range are limited (but see Wright Nelson, 2016).

In the BCCH-CACH contact zone, there is considerable breakdown in the distinctiveness present in the two species' songs in allopatry. Individual birds often produce song with typical characteristics of both species and atypical characteristics of either species in populations near the range interface (Robbins et al., 1986; Ward & Ward, 1974; Curry et al., 2007). The current consensus is that this breakdown is a product of cultural rather than genetic admixture; previous research on song in the BCCH/CACH hybrid zone has found no significant association between

genetic ancestry and song type (Robbins et al., 1986; Sattler et al., 2007; Abbrescia, 2021). This is not surprising given that chickadees are song-learning oscine passerines; if a young chickadee's auditory input includes BCCH-typical, CACH-typical, and intermediate song, it will likely be neurologically capable of learning these songs, and biomechanically capable of producing some version of them later in life. Interestingly, when reared together in a laboratory setting and tape-tutored with both conspecific and heterospecific song, BCCH and CACH developed species-atypical songs that converged upon each other's vocal output rather than copying the species-typical output of the tape tutor, exemplifying the importance of auditory input from neighboring birds to the development of song (Kroodsma & Albano, 1995). Despite their general propensity for learning, the strict maintenance of acoustically-distinct species-typical song outside of the BCCH/CACH hybrid zone suggests an innate disposition to produce song with conspecific acoustic qualities. Any such disposition is likely expressed subtly and is seemingly easily masked by a co-occurring tendency towards copying the local song culture, especially given the apparent lack of association between genetic ancestry and song type found by the studies previously mentioned. These studies focused largely on classification of song in the hybrid zone into BCCH and CACH categories, based on either characteristics of song known to vary dramatically in allopatry (Robbins et al., 1986; Abbrescia, 2021) or by obtaining an axis of largest variation obtained from a multivariate analysis of acoustic measurements (Sattler et al., 2007), which is likely functionally very similar to the former method. While effective in detecting whether individuals can broadly copy BCCH-typical and CACH-typical song, these methods are likely unable to capture more subtle variations in song. Acoustic sampling is another potential limiting factor in previous studies. Either individual chickadees were not recorded for long enough to confidently capture all of their potential song variation (Robbins et al., 1986;

Sattler et al., 2007) or were recorded passively using automated recording units (ARUs) (Abbrescia, 2021), which can neither guarantee individual-level identification of singing birds nor capture high volumes of recordings of the quality needed to take fine-scale acoustic measurements.

A little-studied quality of chickadee song likely to be somewhat genetically constrained is the amount of spectrotemporal variability in song production, which, as previously mentioned, differs substantially in allopatric populations of BCCH (invariable) and CACH (variable). A lab tape-tutoring study of BCCH and CACH showed that CACH have more flexible song learning templates than BCCH, and proposed this as a causal mechanism for their more variable singing tendencies in nature (Wright Nelson, 2016). Similarly, in a population with substantial genetic admixture but a predominantly BCCH-like song culture, the only birds to produce non-BCCH-like song were pure CACH (Abbrescia, 2021). Variability in song production is often investigated by quantifying birds' song repertoires, or the number of different categorical song types they produce (MacDougall-Shackleton, 1997; Kroodsma et al., 1985). For this particular study system, in which the participating species have relatively simple songs, the traditional method of creating song type categories to quantify individual-level song variety is potentially problematic. It is difficult to subjectively delineate song types of structurally simple songs in a way that is likely to be biologically meaningful, since the perceptible units of variation (notes) are so few, and because chickadees in the hybrid zone tend to arrange these few variable units in a large number of different ways (Figure 2.2). There could also be biologically meaningful variation within these units (notes), but this variation will likely be lost in the categorization of songs.

Here I propose alternative methods for quantifying song to address the question of whether genetic ancestry constrains some aspect of otherwise learned song in the BCCH/CACH hybrid zone. I explore the feasibility of active high-volume individual-level recording of chickadee song and whether this approach allows for sufficient sample sizes of individual birds. With these data, I use a multivariate approach to generate acoustic spaces representing the song culture of a hybrid zone chickadee population and quantify the amount of these acoustic spaces occupied by individual birds' songs. This measurement, which I name song variety, serves as a proxy for the relative acoustic variation birds produce in their songs. I then test how well genetic ancestry (quantified in Chapter 1) predicts song variety for individuals in the hybrid zone. Due to the two species' relative levels of song variety in allopatry, I predict a positive relationship between CACH ancestry and song variety.
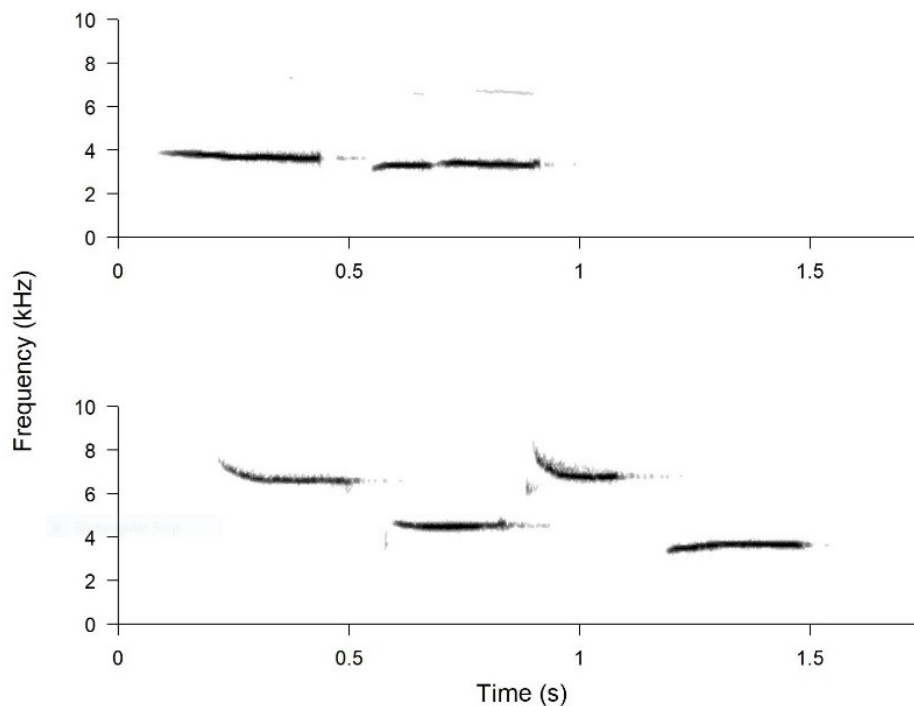
Figure 2.1. Spectrographic depictions of species-typical BCCH (top) and CACH (bottom) songs from Johnson County, MO and Benton County, MO, respectively.

Methods

Song Recording. I recorded songs of 10 color-banded chickadees from which blood samples had been collected at Sparrowfoot Park, Henry County, Missouri. I recorded songs from March to June, 2022, and February to April, 2023, variably during the hours of 06:00 to 16:00.
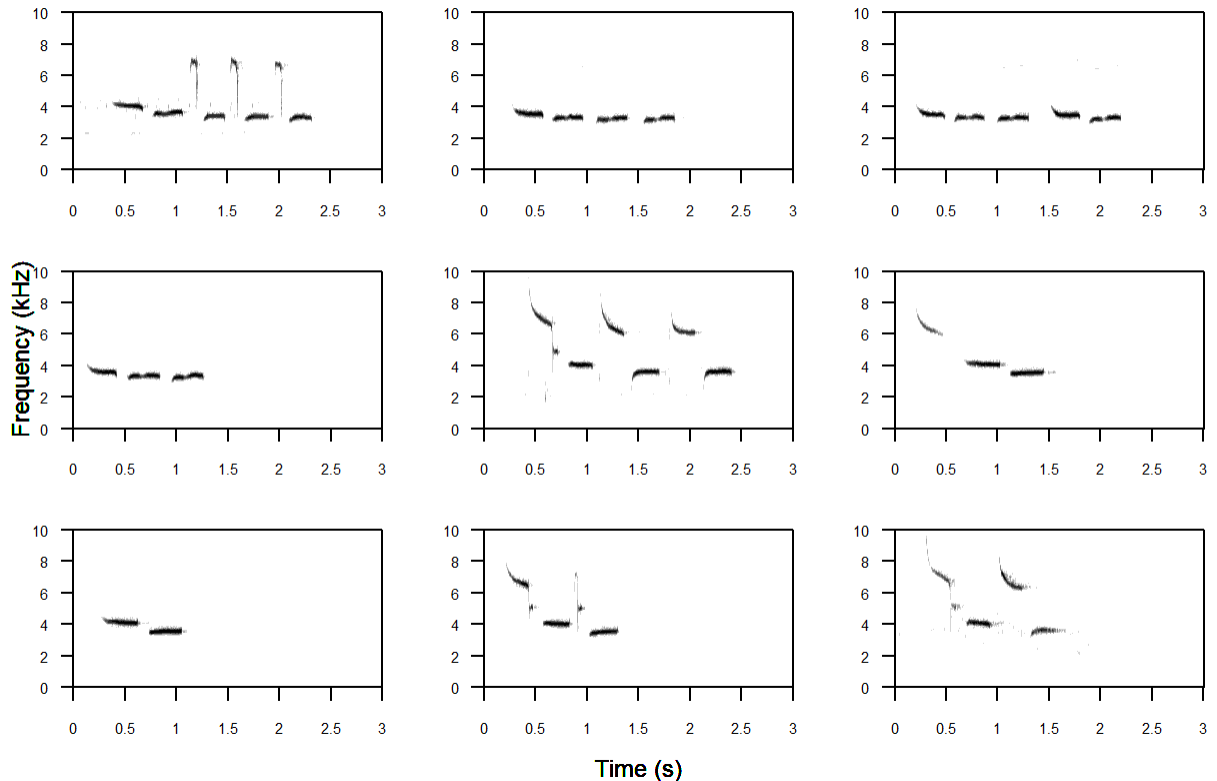
Figure 2.2. Spectrographic depictions displaying the song variety of a chickadee (HZ16) from the hybrid zone population at Sparrowfoot Park, Henry County, MO.

Recordings were made with a Sound Devices MixPre 3 II Audio Recorder/Sennheiser ME 62 microphone mounted in a parabolic dish and stored as 24-bit .wav files with a sampling rate of 48000 Hz.

Firstly, I attempted to re-locate previously captured singing individuals by visiting capture sites and listening for song. If no chickadees were singing upon arrival, I broadcasted BCCH or CACH song and calls from the Merlin app (Cornell Lab of Ornithology) for 1-2

minutes, or until a chickadee in the area responded with song. I then located the singing individual with binoculars to determine whether it was color-banded. If the singing individual was not banded, I recorded its coordinates and continued searching for banded birds. If the singing individual was banded, I immediately began recording its song, getting as close as possible to the individual without disturbing it to ensure both correct color band reading and high recording quality. I then targeted that specific individual for several hours, following it and recording it whenever it sang. If I lost track of the target bird, I used playback of chickadee song and calls to prompt singing and assist in re-location. With the exception of the first week of recording, during which I used song playback from the Merlin app, an audio track composed of songs recorded in Sparrowfoot Park was used to prevent influence of non-local song on the singing behavior of Sparrowfoot Park chickadees. When non-target chickadees responded vocally to the target bird, I quietly dictated indicators to help with identification of the target bird's song in downstream data extraction (e.g., "The louder, 3-note song is the target bird"). Similarly, when song playback was used during recording, I verbally indicated when playback started and ended. Immediately after field recording sessions, I reviewed recordings both visually and aurally in the sound analysis software Raven Pro v.1.6 (K. Lisa Yang Center for Conservation Bioacoustics, 2023) and counted the number of high-quality song iterations (i.e., those with a satisfactory signal-noise ratio and no interference from other vocalizing individuals). To be considered fully recorded, an individual had to (1) exceed 100 high-quality songs' worth of recording, and (2) have been recorded on at least 2 different days. The purpose of these criteria was to prevent excluding some of the variety in an individual's song repertoire due to either limited sampling of song iterations or limited knowledge of the vocal behavior of

the individual, thereby increasing confidence that the full repertoire of each individual had been captured.

Acoustic Data Extraction. Prior to processing, copies of each recording were made and given file names that included relevant metadata (genus, date, color band combination abbreviation, sampling location, recordist initials). Recordings were again reviewed visually and aurally in Raven Pro (Cornell Lab of Ornithology). Because I planned to measure only dominant frequency, extraneous noise that fell within the temporal, but not frequency, range of otherwise high-quality songs was removed with Raven Pro's point-and-click bandpass filtering tool. Songs were then segmented and saved as .wav files that included a single song iteration apiece (1346 songs across 10 individuals). All subsequent processing and data extraction was done in R (v.4.3.1, R Core Team, 2023) using the package seewave (Sueur et al., 2008). The function timer in seewave extracts signal and pause duration measurements from the amplitude profile of a .wav file given a minimum amplitude threshold for signal detection (in % of the maximum amplitude value) and a minimum signal duration (in seconds). As a primary quality control step to check the effectiveness of timer to accurately extract signal periods, I generated spectrograms of each segmented song recording overlaid with detected signal periods given by timer with an amplitude threshold of 10% and a minimum signal duration of 0.05 s. (Figure 2.3). When raising or lowering the amplitude threshold would allow for improved detection of signal periods or elimination of noise, I recorded an adjusted threshold value in a .csv spreadsheet containing file names for each segmented song recording (31% of recordings). When timer captured noise before the song began or after it ended, I recorded approximate start and end times at which to trim the recordings using the function cutw so that only signal periods would be measured (2.1% of recordings). When it was clear that recordings were of too low quality to distinguish signal

29

from noise by amplitude difference, they were marked as such in the .csv spreadsheet and

eliminated from analyses (8.0% of recordings). The parameters designated in this quality control

spreadsheet were written into subsequent data extraction scripts.

I wrote functions to extract the following measurements from each signal period (i.e.,

note) within a song recording: maximum dominant frequency, minimum dominant frequency,

Poecile.sp_Om.YR_Apr182022_SparrowfootPUA.HenryCO.MO_SMP_a_0.45.wav
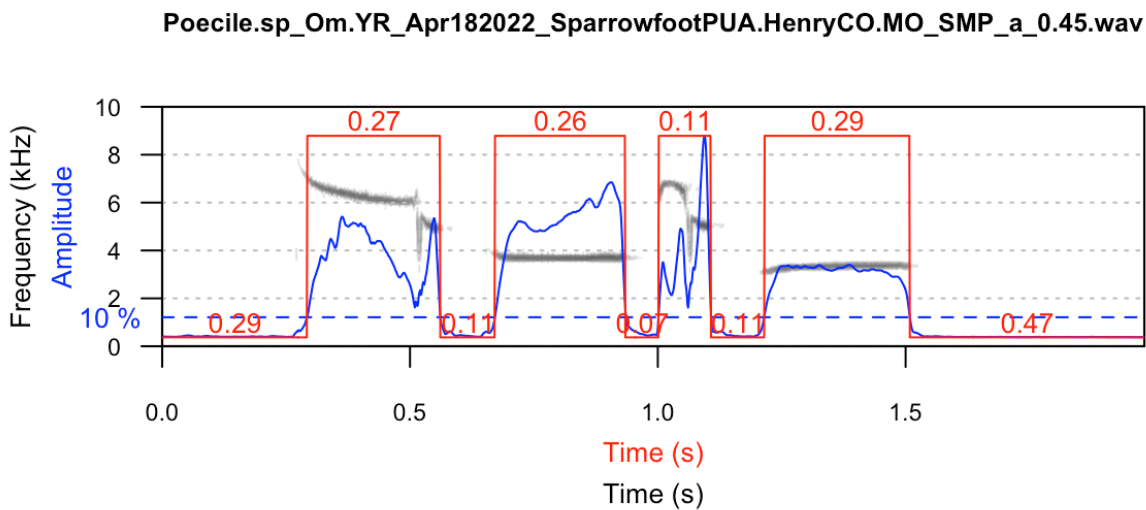
Figure 2.3. Example of a figure generated to check the ability of the seewave (Sueur et al., 2008) function timer to capture signal periods. Spectrographic parameters and labels are displayed in black-and-white, overlaid by the amplitude profiles (blue) and signal/pause durations (red) generated by timer.

mean dominant frequency, median dominant frequency, dominant frequency standard deviation,

maximum absolute dominant frequency slope, and duration. The functions delineated signal

periods using timer with a minimum signal duration of 0.05 s and an amplitude threshold of 10%

unless otherwise noted in the quality control spreadsheet. These note-level measurements were

then used to generate the following song-level measurements for each segmented recording:

number of notes, duration, maximum note duration, minimum note duration, mean note duration,

standard deviation of note duration, signal-pause ratio, maximum dominant frequency, minimum

dominant frequency, standard deviation of note maximum frequency, absolute maximum

30

dominant frequency slope, and standard deviation of absolute maximum dominant frequency slope. These song-level measurements were used in all subsequent analysis steps.

Quantification of Song Variety. To generate a multivariate space representative of the song variety of the Sparrowfoot chickadee population, I performed a principal component analysis (PCA) on scaled and centered values of each song-level variable, excluding signal-noise ratio. To represent the relative amount of this acoustic space occupied by each individual, I generated four minimum spanning trees of a random sample of 100 of its PC scores for four different combinations of the first four principal components. A minimum spanning tree connects each vertex in a set of points to make a tree with the smallest possible edge weight sum (i.e., the least possible Euclidean distance between points). Sums of minimum spanning tree edge weights can be used to quantify the amount of space occupied by points in an *n*-dimensional space (March et al., 2010). This goal could be accomplished alternatively by measuring the two-dimensional area of the minimum convex polygon encompassing a set of points, or by generating clusters of points using a clustering algorithm and counting the number of clusters in which a set of points is represented. However, minimum convex polygon areas tend to over-estimate occupancy for spatially heterogenous data by quantifying empty space and can only be applied to two dimensions, and clustering algorithms are at risk of introducing subjective bias to results due to the need for *a priori* designation of clustering parameters. I found that minimum spanning trees effectively captured the variation in individuals' PC scores while avoiding the pitfalls of other space occupancy quantification methods.

I used the R package PhenotypeSpace (Araya-Salas & Odom, 2022) to generate minimum spanning trees of a random sample of 100 of each individuals' PC scores in the 2-dimensional acoustic space generated by PC1 and PC2. I also generated minimum spanning tree

of individuals' PC scores in the 3-dimensional acoustic space generated by PC1, PC2, and PC3 with the R package emstreeR (Quadros, 2022). I calculated the edge weight sums of each of an individual's minimum spanning trees to quantify the variety of their songs in that space. Lastly, I used PC1 to generate a 1-dimensional measurement of song, a method more representative of what has been done in previous studies of song in the chickadee hybrid zone (notably Sattler et al., 2007). Specifically, I calculated the range of a random sample of 100 of each individual's PC scores on PC1, the axis explaining the most variation in the total dataset of songs.

To ensure that the random sampling of 100 PC scores did not significantly affect the minimum spanning tree edge weight sums, I repeatedly generated 30 minimum spanning trees from random samples of 100 of each individual's PC1 and PC2 scores and checked the standard deviations of each individual's edge weight sums. As expected, edge weight sum standard deviations covaried with the total number of songs recorded, but were generally small (between +/- 0 and +/- 1.37) (Table 2.1). Thus, I concluded that repeated random sampling was not necessary to arrive at accurate individual-level song variety measurements.

Statistical Analysis. To test how well song variety is predicted by genotype, I ran two sets of linear models for each of the 1-, 2-, and 3-dimensional song variety measurements: one with scaled song variety values as the response variable and probability of assignment to CACH + number of songs recorded as predictors, and one with scaled song variety values as the response variable and probability of assignment to CACH as the sole predictor. For each set of models, I then used Akaike information criterion (AIC) to test which of the two models best fit the data. In every case, the model including probability of assignment to CACH as the only predictor had the lower AIC score, signifying a better fit.

Results

PCA of Total Dataset. PC1 explains almost half (46%) of the variation in the total

dataset of songs (Appendix E). An examination of the variable loadings onto PC1 revealed that a

substantial portion of song-level variation in the dataset is credited to a division between songs

with typical BCCH-like and CACH-like characteristics, much like in Sattler et al. 2007.

Table 2.1. Summary statistics of edge weight sums of 30 minimum spanning trees (MST)
generated with 100 randomly-sampled PC1 and PC2 scores for each individual.

Individual ID	Mean MST edge weight sum	Minimum MST edge weight sum	Maximum MST edge weight sum	Standard deviation MST edge weight sum (+/-)	Number of PC scores in full dataset
HZ24	17.09	15.97	17.99	0.558	125
HZ28	30.41	30.41	30.41	0	100
HZ22	36.86	35	38.5	0.9538	125
HZ36	29.11	29.11	29.11	0	100
HZ16	34.69	31.13	37.28	1.374	182
HZ25	20.55	17.08	21.56	0.909	121
HZ3	24.52	22.1	25.07	0.6519	105
HZ30	35.9	31.57	37.1	1.131	108
HZ20	32.52	30.77	35.17	1.078	162
HZ27	27.16	26.08	27.88	0.4046	107

Measurements typically associated with variation between the two species' songs, such as

maximum frequency and note number, weigh heavily onto PC1 in the expected directions given

the typical structure of BCCH and CACH song (Appendix F, Figs. 2.1 & 2.4). PC1 is the only

axis for which the variation explained is clearly interpretable in the context of the study species;

the other two PCs used in analyses, PC2 and PC3, likely explain some of the variation in the dataset that is not easily attributed to one species-specific song type or the other.

Quantification of Song Variety. There was substantial variation in the song variation scores across the sampled individuals in each of the acoustic spaces (Table 2.2, Figs. 2.5-2.6). The quantified differences between song variation scores were in general agreement with the
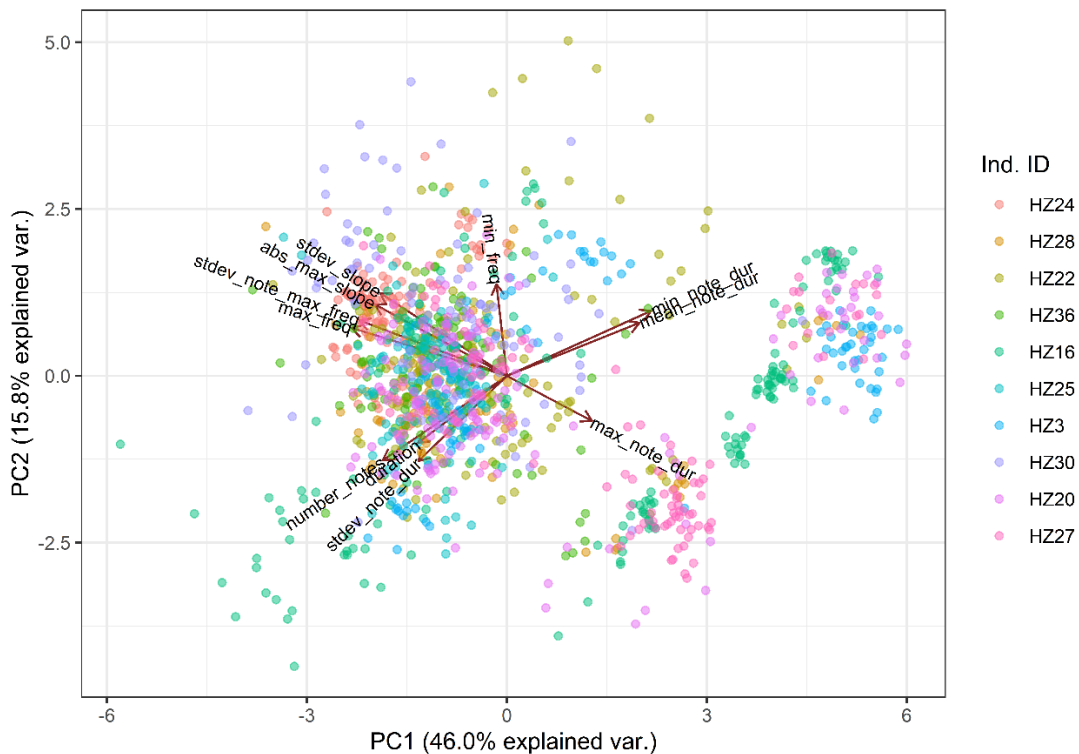
Figure 2.4. A plot of the first two principal components of a PCA of the full dataset of songs (n = 1235) from 10 chickadees in Sparrowfoot Park, Henry County, Missouri. Arrows indicate loading magnitudes and directions of variables.

degree of variation of individual chickadees' song production I perceived while collecting the song data, i.e., birds with more apparent song variety had larger minimum spanning trees, and thus greater edge weight sums.

Linear Regression with Ancestry. For the 2-dimensional and 3-dimensional acoustic spaces created with minimum spanning trees, linear regression showed a positive, very similar

34

relationship between song variety and probability of assignment to CACH with relatively poor fit of the data to the regression line (2-D: $R^2 = 0.1121$, slope = 0.45; 3-D: $R^2 = 0.1015$, slope = 0.44). Between these two spaces, the song variation-ancestry trends among individuals remained fairly constant (Figure 2.7). The 1-dimensional song variety measurement of PC1 score ranges showed a barely positive relationship between song variety and probability of assignment to CACH, with extremely poor fit of the data to the regression line ($R^2 = 0.003$, slope = 0.1) (Figure 2.7).

Table 2.2. Scaled values of song variety measurements in 1, 2 and 3 dimensions.

	PC1 range	PC1-2 minimum spanning tree	PC1-3 minimum spanning tree
HZ24	0.3343	0.5485	0.6327
HZ28	1.098	0.9762	0.9761
HZ22	0.9853	1.183	1.171
HZ36	0.7339	0.9345	1.04
HZ16	1.4	1.114	0.9601
HZ25	0.4975	0.6595	0.6322
HZ3	0.979	0.7872	0.7197
HZ30	1.023	1.152	1.225
HZ20	1.023	1.044	1.098
HZ27	0.956	0.8719	0.8046
Mean	0.903	0.927	0.926
Standard Deviation (+/-)	0.307	0.212	0.217
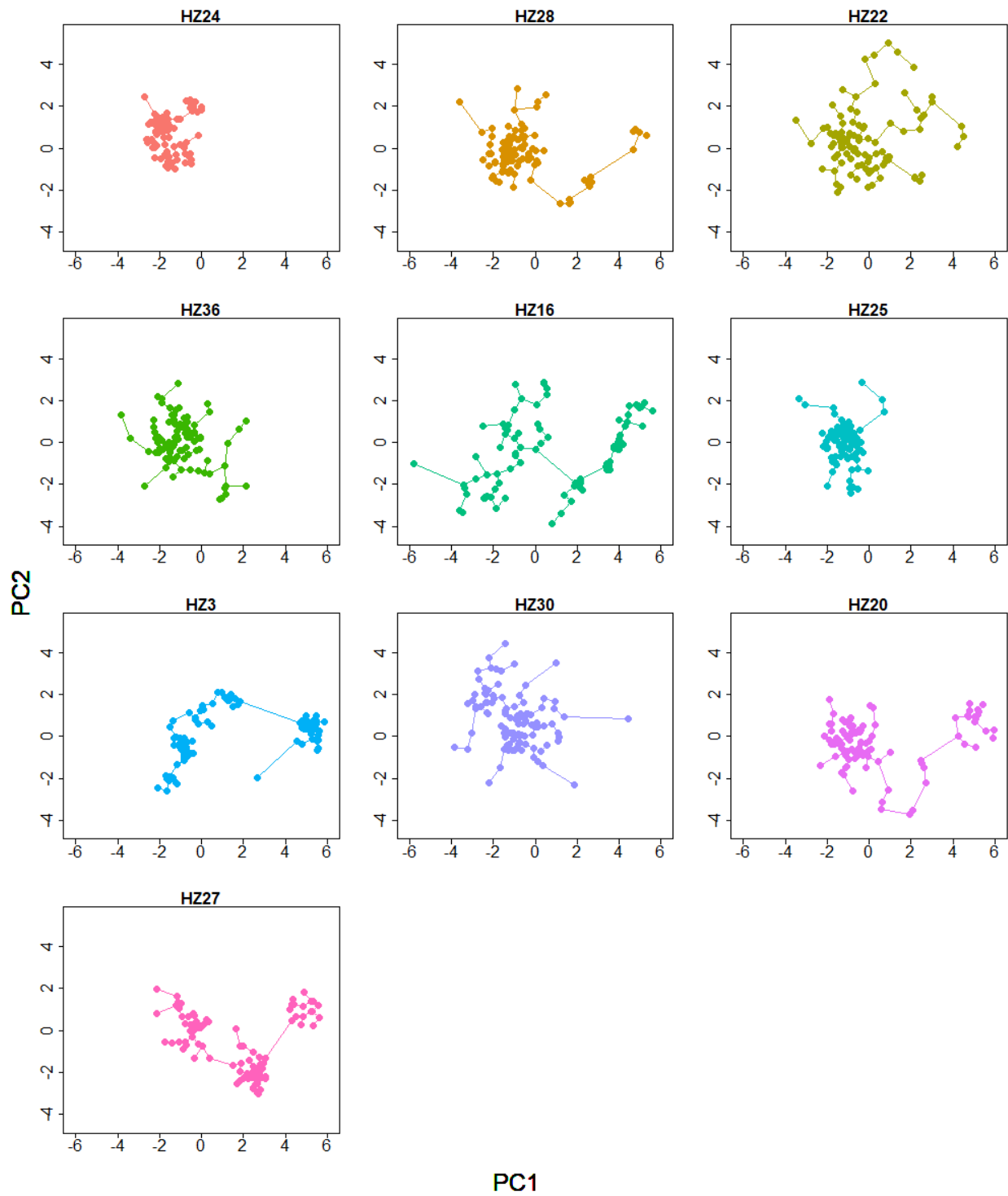
Figure 2.5. Plots of two-dimensional minimum spanning trees for each individual calculated from PC1 and PC2 scores.
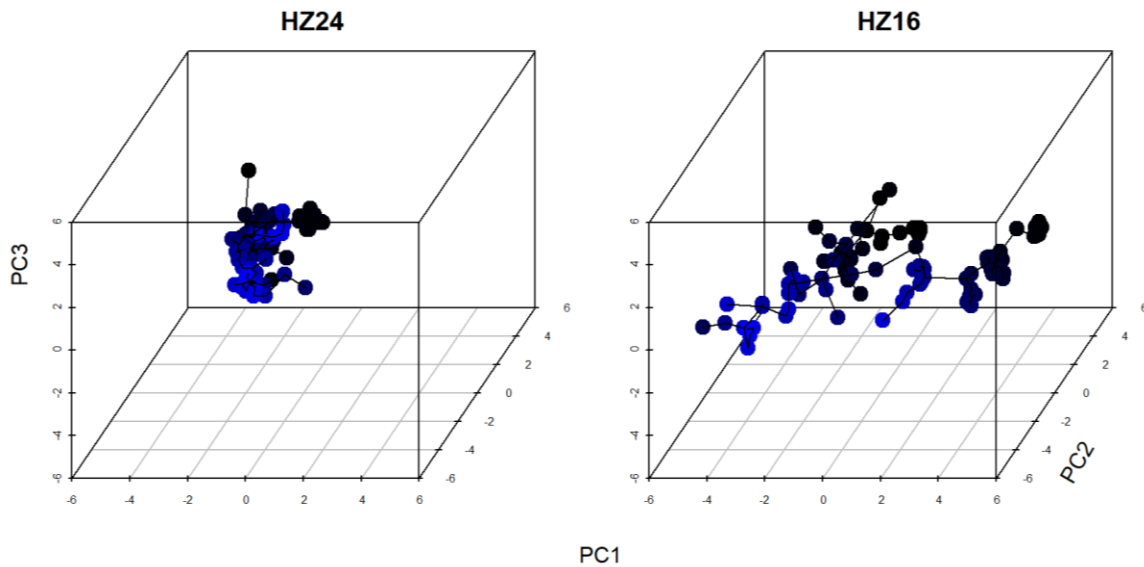
Figure 2.6. Plots of three-dimensional minimum spanning trees calculated from PC1, PC2, and PC3 scores for an individual with less variable song (HZ24) and an individual with highly variable song (HZ16). The spread of points across PC2 is represented by point color (amount of blue decreases with increasing PC2 score).

Discussion

Previous studies on the relationship between song and ancestry in the BCCH/CACH hybrid zone have prioritized sampling of individuals over sampling of songs. In order to properly perform song variety analyses, I needed to deviate from this route and actively obtain high volumes of high-quality recordings from individuals, something that has not been attempted in this hybrid zone to my knowledge. To control for the potentially confounding effect of microdialectal variation due to cultural isolation, I also needed to limit sampling to a single population without physical barriers to the cultural transmission of song. I was able to record 100 or more songs apiece from individual color-banded and genotyped chickadees in Sparrowfoot Park over the course of ~1.5 standard field seasons, demonstrating that this approach to acoustic sampling is possible, albeit challenging. The tradeoff between song sampling and individual sampling was significant: in the time I was able to carry out my field work, I fully recorded only
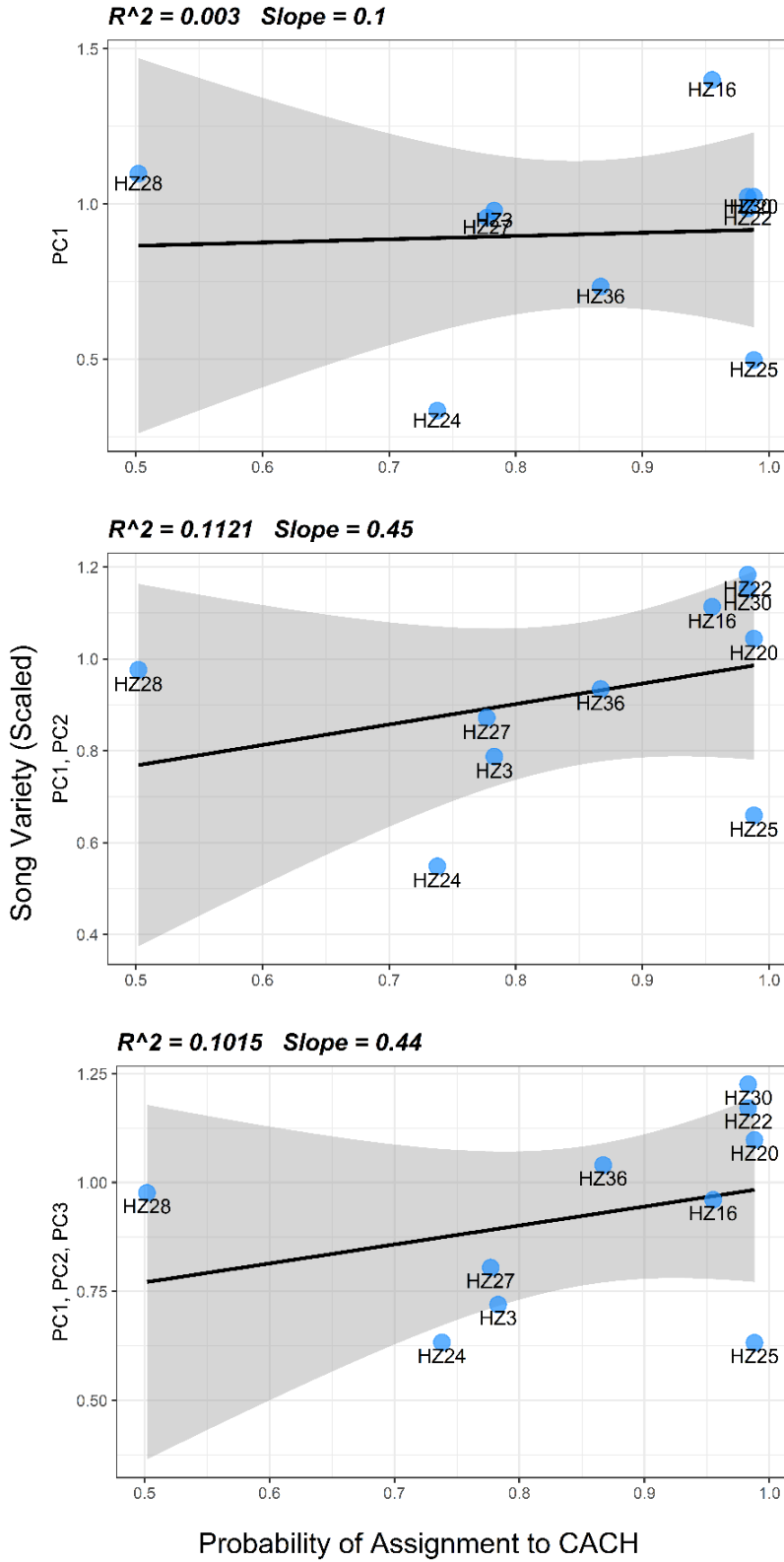
Figure 2.7. Linear regressions modeling the relationship between chickadee ancestry (via probability of assignment to CACH) and scaled song variety values in 1-dimensional, 2-dimensional, and 3-dimensional acoustic spaces generated with PCs 1-3.

10 chickadees. The low sample size of individuals and the decision to sample within only one population to limit microdialectal variation also impacted my sampled genotype diversity. Of the 10 fully-recorded chickadees, none had predominantly BCCH ancestry, and only one appeared to be a first-generation hybrid; the rest had all or predominantly CACH ancestry. The limitations of this approach to acoustic sampling must, however, be considered alongside the short timeframe in which this study was conducted. The implementation of these sampling methods over a longer time period and with more researchers in the field would likely be able to produce sample sizes large enough for conclusive song-genotype analyses.

Quantifying individual-level variation in birdsong without categorizing song types is, to my knowledge, a yet-unexplored avenue of acoustic analysis. The song variety scores obtained from the minimum spanning tree edge weights were variable among individuals, and this variation reflected the levels of variation among individuals I perceived during extensive time spent recording them in the field. I used R for all acoustic measurement and analysis steps both for time efficiency and reproducibility of my exploratory song quantification methods. Future studies can replicate each step of this process, from recording quality control to automated acoustic measurement extraction to generation of multivariate acoustic space and song variety scores using the R scripts I wrote for these purposes.

None of the linear models of the relationship between ancestry and song variety showed a degree of fit to the data that would allow for confident assertion that this relationship is unambiguous. However, some takeaways from these models provide support for the potential of multidimensional song variety to reveal biologically important patterns. Interpretation of the PCA used to create song variety scores suggests that PC1, the axis of greatest variation in the data, is responsible for the division between songs with BCCH-typical and CACH-typical

acoustic qualities. This exemplifies the degree of influence of the two species' distinct song types on the song culture of an admixed population. However, of the three linear models predicting song variety with ancestry, the model with only the range in PC1 as the response variable stood out as the worst-fitting to the data. This result is in line with past research on song in the BCCH/CACH hybrid zone finding no relationship between genotype and categorical song type (Robbins et al., 1986; Sattler et al., 2007; Abbrescia, 2021). Besides producing better-fitting models with ancestry, the 2-dimensional and 3-dimensional song variety scores showed very similar values of the predicted positive relationship between song variety and ancestry (regression line slope of 0.45 and 0.44, respectively), and better maintained the relative positions of ancestry vs. song variety among individuals. These patterns were present and consistent despite the aforementioned limitations of small sample size of individuals and limited genotype variety. Taken together, these interpretations suggest that other dimensions of variation besides the broad split of BCCH-like and CACH-like song may be related to ancestry, and worth further exploring.

The results of this study provide evidence for the usefulness of multidimensional song variety scores for honestly quantifying vocal variation in songbirds and justify their continued use in avian bioacoustics on multiple fronts. Firstly, with properly-directed time and effort in the field, researchers can obtain sufficient volumes of high-quality recordings to conduct these analyses. Secondly, the R scripts I wrote to carry out this study will offset the time-demanding process of acoustic measurement extraction and analysis for future researchers, and can be easily tailored to other songbird taxa with different acoustic song structure. Thirdly, multidimensional song variety measurements showed a stronger (albeit still inconclusive) relationship to ancestry than did a 1-dimensional song variety measurement of range along the greatest axis of variation

in the data, suggesting the importance of multidimensional song quantification in taxa whose songs are influenced by cultural evolution. Lastly, and perhaps most importantly, measuring song variety in this manner removes the need for subjective categorization of song types. This is a crucial advancement especially for taxa whose songs are not easily delineated into obviously discrete categories, as is the case in the BCCH/CACH hybrid zone. For studies of vocal variety, the methods I develop here have the potential to uncover subtle, but biologically important variation that is likely lost in the quantification of overly-broad categories of vocalizations. Teasing out this hidden variation could enhance our understanding of what drives song development and transmission in vocal learning taxa. I advocate for the continued use and fine-tuning of these methods in research on vocal variety in songbirds, and hopefully their eventual expansion to other areas of bioacoustics.

References

Abbrescia, A. (2021). Dawn chorus male song patterns in relation to ancestry in the Black-capped Chickadee × Carolina Chickadee hybrid zone. *Villanova Research Journal*, *3*, 47-57.

Araya-Salas, M., & Odom, K. (2022). *PhenotypeSpace: an R package to quantify and compare phenotypic trait spaces*. R package version 0.1.0.

Beecher, M., & Brenowitz, E. (2005). Functional aspects of song learning in songbirds. *Trends in Ecology & Evolution*, *20*(3), 143–149.

Catchpole, C. K., & Slater, P. J. B. (2008). *Bird song: Biological themes and variations, Second edition*. Cambridge University Press, New York.

Christie, P. J., Mennill, D. J., & Ratcliffe, L. M. (2004). Pitch shifts and song structure indicate male quality in the dawn chorus of black-capped chickadees. *Behavioral Ecology and Sociobiology*, *55*(4), 341–348.

Curry, R. L., Rossano, L. M., & Reudink, M. W. (2007). Behavioral aspects of chickadee hybridization. In *The ecology and behavior of chickadees and titmice: An integrated approach*. Oxford University Press.

DeRaad, D. A., Applewhite, E. E., Tsai, W. L. E., Terrill, R. S., Kingston, S. E., Braun, M. J., & McCormack, J. E. (2023). Hybrid zone or hybrid lineage: A genomic reevaluation of Sibley's classic species conundrum in *Pipilo* towhees. *Evolution*, *77*(3), 852–869.

Goller, F. (2021). Vocal athletics – from birdsong production mechanisms to sexy songs. *Animal Behaviour*, S0003347221001081.

Kelemen, E. P., Zusi, K. E., & Curry, R. L. (2015). Song repertoire of Carolina Chickadees (*Poecile carolinensis*) in southeastern Pennsylvania. *The Wilson Journal of Ornithology*, *127*(2), 271–276.

K. Lisa Yang Center for Conservation Bioacoustics at the Cornell Lab of Ornithology. (2023). Raven Pro: Interactive Sound Analysis Software (Version 1.6.4) [Computer software]. Ithaca, NY: The Cornell Lab of Ornithology.

Kroodsma, D. E., & Canady, R. A. (1985). Differences in Repertoire Size, Singing Behavior, and Associated Neuroanatomy Among Marsh Wren Populations Have a Genetic Basis. *The Auk*, *102*(3), 439–446.

Kroodsma, D. E., & Konishi, M. (1991). A suboscine bird (eastern phoebe, Sayornis phoebe) develops normal song without auditory feedback. *Animal Behaviour*, *42*(3), 477–487.

Kroodsma, D. E., & Albano, D. J. (1995). Song development by Black-capped Chickadees (*Parus atricapillus*) and Carolina Chickadees (*P. carolinensis*). *The Auk*, *112*(1), 29–43.

Kroodsma, D. E., Byers, B. E., Hill, C., Farrington, J., Gill, F. B., Stoddard, P. K., & Wilda, K. (1999). Geographic variation in Black-capped Chickadee songs and singing behavior. *The Auk*, *116*(2), 387–402.

Macdougall-Shackleton, S. A. (1997). Sexual selection and the evolution of song repertoires. In V. Nolan, E. D. Ketterson, & C. F. Thompson (Eds.), *Current Ornithology* (pp. 81–124). Boston, MA: Springer US.

March, W. B., Ram, P., & Gray, A. G. (2010). *Fast Euclidean Minimum Spanning Tree: Algorithm, Analysis, and Applications*.

Marler, P. (1990). Song learning: The interface between behaviour and neuroethology. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *329*(1253), 109–114.

Mason, N. A., Burns, K. J., Tobias, J. A., Claramunt, S., Seddon, N., & Derryberry, E. P. (2017). Song evolution, speciation, and vocal learning in passerine birds. *Evolution*, *71*(3), 786–796.

Nowicki, S., & Searcy, W. A. (2014). The evolution of vocal learning. *Current Opinion in Neurobiology*, *28*, 48–53.

Quadros, A. (2022). *emstreeR: Tools for Fast Computing and Plotting Euclidean Minimum Spanning Trees*. R package version 3.0.0.

Robbins, M. B., Braun, M. J., & Tobey, E. A. (1986). Morphological and vocal variation across a contact zone between the chickadees Parus atricapillus and P. carolinensis. *The Auk*, *103*(4), 655–666.

Sattler, G. D., Sawaya, P., & Braun, M. J. (2007). An assessment of song admixture as an indicator of hybridization in Black-capped Chickadees (*Poecile atricapillus*) and Carolina chickadees (*P. carolinensis*). *The Auk, 124*(3): 926–944.

Soha, J. A., & Marler, P. (2000). A species-specific acoustic cue for selective song learning in the white-crowned sparrow. *Animal Behaviour*, *60*(3), 297–306.

Ward, R. (1966). Regional variation in the song of the Carolina chickadee. *The Living Bird*, *5*, 127-150.

Ward, R., & Ward, D. A. (1974). Songs in contiguous populations of black-capped and Carolina chickadees in Pennsylvania. *The Wilson Bulletin, 86*(4), 344–356.

Wright Nelson, S. G. (2016). Song learning, song variation, and cultural change in two hybridizing sister songbird species, Black-capped (Poecile atricapillus) and Carolina (P. carolinensis) Chickadees. The Ohio State University, Columbus, Ohio.

Sueur, J., Aubin, T., & Simonis, C. (2008). Seewave: a free modular tool for sound analysis and synthesis. *Bioacoustics*, *18*, 213-226.

SUMMARY

In Chapter 1, I applied 9 autosomal SNP-RFLP molecular markers designed to genotype BCCH, CACH, and their hybrids to two putative hybrid zone populations and known pure-species chickadee populations in western Missouri and eastern Kansas. I found predictable levels of genetic admixture in both putative hybrid zone populations, as well as low levels of genetic admixture in pure-species populations in western Missouri. Ruling out underestimation of hybrid zone width as a cause of this unexpected admixture, I concluded that the markers, while seemingly species-diagnostic in the eastern portion of the hybrid zone, are likely not fixed in their respective species' ranges in Missouri. I also tested for differences in levels of linkage disequilibrium (LD) between physically-linked and physically-unlinked groups of the markers. While one physically linked marker pair showed high values of LD, overall levels of LD did not appear to be disproportionately driven by physical linkage.

In Chapter 2, I tested the feasibility of obtaining high volumes of high-quality individual-level song recordings of chickadees in a BCCH/CACH hybrid zone population. I found that, while possible, taking this approach to acoustic sampling requires heavy time investment in data collection in order to also obtain the sample size of individuals needed to conduct conclusive analyses. To these data I applied a novel method of quantifying individual-level song variation by reducing the dimensionality of a dataset of acoustic measurements using principal component analysis (PCA) and generating minimum spanning trees of equal subsets of individuals' PC scores in the 2-dimensional and 3-dimensional spaces produced by the three PCs explaining the most variation in the dataset. I also generated a 1-dimensional measurement of song variation using only PC1, which explained almost half of the variation in the total dataset, to represent a

more traditional method of quantifying song. I found that the measurements of song variety showed appreciable variation among individuals that was concordant with the individual-level song variety I observed while recording them in the field. I ran linear models to assess how well ancestry predicted 1-dimensional, 2-dimensional, and 3-dimensional song variety measurements. While none of the models showed a tight fit to the data, the models with the 2-dimensional and 3-dimensional song variety measurements upheld my prediction of a positive relationship between CACH ancestry and song variety. The model with the 1-dimensional song variety measurement showed a notably worse fit to the data than the models with the multidimensional measurements and was barely positive. Although not a conclusive result, the appearance of these trends despite low sample size and low genotype variety suggests that multidimensional song variation could be an overlooked characteristic of song with biological importance in BCCH and CACH.

The results of this study are in line with others that have addressed the question of the relationship between ancestry and song in the hybrid zone: song learning and assimilation to the local song culture is clearly prevalent, and I was unable to demonstrate a strong relationship between genotype and song variety. However, the weak signal of increasing song variety with increasing CACH ancestry, together with the notable difference in song variety between allopatric BCCH and CACH and the previously mentioned evidence for more flexible song learning (Wright-Nelson, 2016) and production (Abbrescia, 2021) in CACH cannot be ignored. In a longer-term study, perhaps 4 or 5 years, implementation of acoustic sampling and analysis methods similar to mine could likely sample enough individuals with enough variety in genetic composition to run these same analyses and obtain a more definite answer to the song variety-genotype question. Preferably, a study over this timeframe would also implement a higher-

resolution genotyping method; output from a reduced-representation sequencing approach such as RAD-Seq could be aligned to the BCCH genome assembly and SNPs found to be highly divergent in parental controls could be used for genotyping, following Alexander et al. 2022. Regardless of the outcome, obtaining a conclusive answer to the song variety-genotype question in the BCCH/CACH hybrid zone will enhance our understanding of the drivers behind chickadee song development and prompt further questions on how best to study the evolution of learned vocalizations.

ADDITIONAL REFERENCES

Alexander, A., Robbins, M. B., Holmes, J., Moyle, R. G., & Peterson, A. T. (2022). Limited movement of an avian hybrid zone in relation to regional variation in magnitude of climate change. *Molecular Ecology*, *31*(24), 6634–6648.

Barton, N. H., Hewitt, G. M. (1989). Adaptation, speciation and hybrid zones. *Nature, 341*(6242), 497–503.

Beecher, M., & Brenowitz, E. (2005). Functional aspects of song learning in songbirds. *Trends in Ecology & Evolution*, *20*(3), 143–149.

Bennett, K. F. P., Lim, H. C., & Braun, M. J. (2021). Sexual Selection and Introgression in Avian Hybrid Zones: Spotlight on Manacus. *Integrative and Comparative Biology*, *61*(4), 1291–1309.

Braun, M. J., & Robbins, M. B. (1986). Extensive Protein Similarity of the Hybridizing Chickadees Parus atricapillus and P. carolinensis. *The Auk*, *103*(4), 667–675.

Bronson, C. L., Grubb, T. C., & Braun, M. J. (2003). A test of the endogenous and exogenous selection hypotheses for the maintenance of a narrow avian hybrid zone. *Evolution*, *57*(3), 630–637.

Bronson, C. L., Grubb, T. C., Sattler, G. D., & Braun, M. J. (2005). Reproductive success across the Black-capped chickadee (Poecile atricapillus) and Carolina chickadee (P. carolinensis) hybrid zone in Ohio. *The Auk*, *122*(3), 759–772.

Curry, R. L., Rossano, L. M., & Reudink, M. W. (2007). Behavioral aspects of chickadee hybridization. In *The ecology and behavior of chickadees and titmice: An integrated approach*. Oxford University Press.

Emlen, S. T., Rising, J. D., & Thompson, W. L. (1975). A behavioral and morphological study of sympatry in the Indigo and Lazuli buntings of the Great Plains. *The Wilson Bulletin*, 87(2), 145–302.

Endler, J. A. (1978). *Geographic variation, speciation, and clines*. Princeton, N.J.: Princeton University Press.

Kenyon, H. L., Toews, D. P. L., & Irwin, D. E. (2011). Can Song Discriminate between MacGillivray's and Mourning Warblers in a Narrow Hybrid Zone? *The Condor*, *113*(3), 655–663.

Kenyon, H. L., Alcaide, M., Toews, D. P. L., & Irwin, D. E. (2017). Cultural isolation is greater than genetic isolation across an avian hybrid zone. *Journal of Evolutionary Biology*, *30*(1), 81–95.

Marler, P. (1990). Song learning: The interface between behaviour and neuroethology. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *329*(1253), 109–114.

McQuillan, M. A., Huynh, A. V., Taylor, S. A., & Rice, A. M. (2017). Development of 10 novel SNP-RFLP markers for quick genotyping within the black-capped (Poecile atricapillus) and Carolina (P. carolinensis) chickadee hybrid zone. *Conservation Genetics Resources*, *9*(2), 261–264.

Ottenburghs, J., Ydenberg, R. C., Van Hooft, P., Van Wieren, S. E., & Prins, H. H. T. (2015). The Avian Hybrids Project: Gathering the scientific literature on avian hybridization. *Ibis*, *157*(4), 892–894.

Reudink, M. W., Mech, S. G., & Curry, R. L. (2006). Extrapair paternity and mate choice in a chickadee hybrid zone. *Behavioral Ecology*, *17*(1), 56–62.

Sattler, G. D., Sawaya, P., & Braun, M. J. (2007). An assessment of song admixture as an indicator of hybridization in Black-capped Chickadees (*Poecile atricapillus*) and Carolina chickadees (*P. carolinensis*). *The Auk, 124*(3): 926–944.

Semenov, G. A., Linck, E., Enbody, E. D., Harris, R. B., Khaydarov, D. R., Alström, P., Andersson, L., & Taylor, S. A. (2021). Asymmetric introgression reveals the genetic architecture of a plumage trait. *Nature Communications*, *12*(1), 1019.

Soha, J. A., & Marler, P. (2000). A species-specific acoustic cue for selective song learning in the white-crowned sparrow. *Animal Behaviour*, *60*(3), 297–306.

Taylor, S. A., White, T. A., Hochachka, W. M., Ferretti, V., Curry, R. L., & Lovette, I. (2014). Climate-Mediated Movement of an Avian Hybrid Zone. *Current Biology*, *24*(6), 671–676.

Taylor, S. A., Curry, R. L., White, T. A., Ferretti, V., & Lovette, I. (2014). Spatiotemporally consistent genomic signatures of reproductive isolation in a moving hybrid zone: Genomics of a moving hybrid zone. *Evolution*, *68*(11), 3066–3081.

Taylor, S. A., Larson, E. L., & Harrison, R. G. (2015). Hybrid zones: Windows on climate change. *Trends in Ecology & Evolution*, *30*(7), 398–406.

Toews, D. P. L., Taylor, S. A., Vallender, R., Brelsford, A., Butcher, B. G., Messer, P. W., & Lovette, I. J. (2016). Plumage Genes and Little Else Distinguish the Genomes of Hybridizing Warblers. *Current Biology*, *26*(17), 2313–2318.

Wheatcroft, D., & Qvarnström, A. (2017). Genetic divergence of early song discrimination between two young songbird species. *Nature Ecology & Evolution*, *1*(7), 0192.

Wright Nelson, S. G. (2016). Song learning, song variation, and cultural change in two hybridizing sister songbird species, Black-capped (Poecile atricapillus) and Carolina (P. carolinensis) Chickadees. The Ohio State University, Columbus, Ohio.

APPENDICES

Appendix A. Research Compliance Certificates

Missouri State
UNIVERSITY

May 1, 2023

RE: IACUC protocol 2021-22

Shleby Palmer,

IACUC protocol #2021-22 entitled "Evolution and ecology of passerine birds" was approved by the committee on January 21, 2022 and expires January 20, 2025.

The protocol reflects that you are approved to work with Dr. Jay McEntee on this project.

Thank you and if you need anything in the future regarding this protocol please contact me either via email (johnnapedersen@missouristate.edu) or at 417-836-3737.

Sincerely,

Johnna Pedersen

Johnna Pedersen
IACUC Administrator/Member
Interim Director of Research Administration

Protocol 2021-22   IACUC approved 1/21/2022 - 1/20/2025

Missouri State
UNIVERSITY

ANIMAL CARE & USE APPLICATION

INSTITUTIONAL ANIMAL CARE & USE COMMITTEE
v. July 2019

All Animal Care & Use Applications should be submitted electronically to IACUC@missouristate.edu.

A. Investigator Information		
Principal Investigator: Jay McEntee	Department: Biology	Office Address: 221 Temple Hall
Office Phone: 5103884065(cell)	Emergency Phone: 5103884065	Email: jaymcentee@missouristate.edu

B. Project Information

Project Title: Evolution and ecology of passerine birds

Protocol Action:	Protocol Type:	Protocol Class:
☒ New Proposal	☒ Research	☐ Agricultural
☐ Pilot Study	☐ Teaching	☒ Behavioral
☐ Renewal (due to protocol expiration)		☐ Biomedical
☐ Review for Exemption		☒ Wildlife/Conservation

Is this project externally funded and/or do you anticipate future funding? ☐ Yes ☒ No
If Yes, what is the name of the Funding Agency and grant number/title?

C. Previous Approved Protocol

For work that is similar to a previously approved protocol, provide the original protocol number and approval date. On the remainder of the forms, indicate changes to the originally approved protocol in bold font.

Original Protocol Number:	Approval Date:
2020-27	January 5, 2021

D. Investigator Assurances

1. The information provided herein is accurate and, to the best of my knowledge, conforms to all applicable University, PHS, and USDA policies on the use of animals in research and teaching.
2. I affirm that all procedures involving vertebrate animals will be performed only by personnel trained in the humane care, handling, and use of animals and that all personnel will abide by the recommendations of the University's Occupational Health and Safety Program.
3. I agree not to proceed with any portion of this project until I receive written approval from the Missouri State University Institutional Animal Care and Use Committee.
4. I agree any changes in the procedures contained in this protocol will be promptly forwarded to the IACUC for review and approval prior to performing any revised procedures.
5. I agree to provide proper, current documentation (e.g., licenses, permits and additional approval forms), when applicable, to the Office of Research Compliance throughout the course of this project.
6. I agree to allow inspection of my research facilities by members of the IACUC, including the veterinarian, and to comply promptly if informed of any violations of the Missouri State University's Animal Care and Use Policy.
7. I have taken into consideration the three "Rs" (replacement, reduction, and refinement) for my study and provided adequate justification for the animal model chosen, animal numbers requested, and procedures to reduce pain and distress.
8. I have conducted a literature search to ensure that I am not unnecessarily duplicating previous experiments.

Jay McEntee	December 1, 2021
Signature of Principal Investigator	Date

50

Appendix B. PCR protocol information from McQuillan et al. 2017.

Locus ID	F Primer (5'-3')	R Primer (5'-3')	Annealing Temp. (ºC)	PCR Product Length (bp)
c0p171	TCACACAACAGCCATTTGTACG	TTCCACATGGGATATCTTGC	56.4	211
c0p183	TCTTCTGCAGGAAGTAAACACTG	AAGTTCAGGACTCCATTTTGCG	57	301
c0p184	ATCCATGCAGAAAGACTGGG	TCACTCTCAGTGCATAAGCTCC	57.7	221
c0p238	TGGAGATGCAGTTTTTCTCCC	AAGAGGAGATGGCTTTTCGC	53.5	267
c0p251	AACCAACAGCTCTGTTCCCTG	AATGCAGAAGGCAGGAACCTC	58	183
c0p283	TTTCTTTCCTGCAAAGCCCC	TAAGGAGGGCCGTCCAACAC	56.4	279
c0p303	TAGAGGCAGCAGTGGTACAG	ATGAGCGGTGTTTGTTCCTTG	55.8	291
c0p356	AGGAATTCCACTGGTAACTTACTG	TGAAGAACAGTCATTAGTTTTGGC	54	200
c0p373	GACTGCAATCCTCTTGGCTG	CAGCCTCAATGCAATCCACAG	55.8	274
c0p628	ACACTCCTGCATCTTCCCTTC	AGGCTTTGGGATTCTTTGTGC	55	217

Appendix C. RFLP protocol information from McQuillan et al. 2017.

Locus ID	Restriction enzyme	Digestion conditions	BCCH allele	CACH allele	BCCH allele fragment length (bp)	CACH allele fragment length (bp)
c0p171	*Alu*I	1-16 hr. at 37ºC; 20 min. at 80ºC	G	A	119; 92	211
c0p183	*Hpy*CH4III	1 hr. at 37ºC; 20 min. at 65ºC	T	G	152; 85; 64	237; 64
c0p184	*Hpy*CH4III	1 hr. at 37ºC; 20 min. at 65ºC	T	G	150; 71	221
c0p238	*Nsp*I	1-16 hr. at 37ºC; 20 min. at 65ºC	T	A	120; 111; 36	231; 36
c0p251	*Bcc*I	1 hr. at 37ºC; 20 min. at 65ºC	C	T	116; 67	183
c0p283	*Nco*I	1-16 hr. at 37ºC; 20 min. at 80ºC	A	T	279	153; 126
c0p303	*Acc*I	1-16 hr. at 37ºC; 20 min. at 80ºC	G	A	150; 141	291
c0p356	*Hpy*188III	1 hr. at 37ºC; 20 min. at 65ºC	G	A	129; 71	200
c0p373	*Hpy*188III	1 hr. at 37ºC; 20 min. at 65ºC	C	T	274	137; 137
c0p628	*Bsa*AI	1-16 hr. at 37ºC	C	T	107; 110	217

Appendix D: Chickadee Tissue Loan Metadata.

KU Accession #	Individual ID	Date	Coordinates	Species	City/Area	County	State
132836	BC10	10/30/15	39.22200, -96.61100	*Poecile atricapillus*	Manhattan	Riley	KS
131638	BC2	2/9/16	39.03200, -95.20600	*Poecile atricapillus*	Lawrence	Douglas	KS
133469	BC11	1/13/15	39.038978, -94.762190	*Poecile atricapillus*	Lake Quivira	Johnson	KS
132058	BC3	4/13/16	38.27940, -94.42730	*Poecile atricapillus*	Butler Lake	Bates	MO
132059	BC4	4/13/16	38.27940, -94.42730	*Poecile atricapillus*	Butler Lake	Bates	MO
132060	BC5	4/13/16	38.28430, -94.43420	*Poecile atricapillus*	Butler Lake	Bates	MO
132061	BC6	4/13/16	38.28090, -94.42790	*Poecile atricapillus*	Butler Lake	Bates	MO
132062	BC7	4/13/16	38.28090, -94.42790	*Poecile atricapillus*	Butler Lake	Bates	MO
132063	BC8	4/13/16	38.28220, -94.43360	*Poecile atricapillus*	Butler Lake	Bates	MO
132064	BC9	4/13/16	38.28220, -94.43360	*Poecile atricapillus*	Butler Lake	Bates	MO
92269	CA2	11/21/01	36.60161, -93.32783	*Poecile carolinensis*	Table Rock Lake	Stone	MO
92270	CA3	11/21/01	36.60161, -93.32783	*Poecile carolinensis*	Table Rock Lake	Stone	MO
132050	CA4	4/4/2016	37.87500, -93.71300	*Poecile carolinensis*	Bird Song CA	St. Clair	MO
132051	CA5	4/4/2016	37.87500, -93.71300	*Poecile carolinensis*	Bird Song CA	St. Clair	MO
132052	CA6	4/4/2016	37.87500, -93.71300	*Poecile carolinensis*	Bird Song CA	St. Clair	MO

132054	CA8	4/4/2016	37.87500, -93.71300	*Poecile carolinensis*	Bird Song CA	St. Clair	MO
132055	CA9	4/4/2016	37.87500, -93.71300	*Poecile carolinensis*	Bird Song CA	St. Clair	MO
132056	CA10	4/4/2016	37.87500, -93.71300	*Poecile carolinensis*	Bird Song CA	St. Clair	MO
132057	CA11	4/4/2016	37.87500, -93.71300	*Poecile carolinensis*	Bird Song CA	St. Clair	MO

Appendix E: Song Measurement PCA Summary

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11
Eigenvalue	2.251	1.318	1.198	1.167	0.8354	0.7015	0.3592	0.1919	0.154	0.1153	0.08926
Proportion of Variance	0.4605	0.1579	0.1305	0.1237	0.06345	0.04473	0.01173	0.00335	0.00215	0.00121	0.00072
Cumulative Proportion of Variance	0.4605	0.6184	0.7489	0.8726	0.9361	0.9808	0.9926	0.9959	0.9981	0.9993	1

Appendix F: Variable Loadings of PC1-PC4

Measurement	PC1	PC2	PC3	PC4
Number of notes	-0.3139	-0.3643	-0.02424	0.3696
Duration	-0.2244	-0.289	-0.2965	0.5202
Maximum note duration	0.2154	-0.1951	-0.5218	-0.3256
Minimum note duration	0.3648	0.2806	-0.1889	0.2945
Mean note duration	0.3356	0.2288	-0.4251	0.06295
Standard deviation of note duration	-0.2242	-0.3674	-0.0387	-0.5856
Maximum frequency	-0.3927	0.2072	-0.0244	0.0631
Minimum frequency	-0.02616	0.3959	0.4537	-0.06543
Standard deviation of note maximum frequency	-0.3731	0.2394	-0.08733	0.06898
Absolute maximum dominant frequency slope	-0.3344	0.3084	-0.3268	-0.1224
Standard deviation of absolute maximum dominant frequency slope	-0.3206	0.3572	-0.3168	-0.1588