Exploration of a Novel Avian Hybrid Zone Along the Front Range of the Colorado Rocky Mountains

William C. Anderson

Department of Ecology and Evolutionary Biology,

University of Colorado at Boulder

Defense Date: March 14th, 2022

Thesis Advisor:
Scott A. Taylor, PhD, Department of Ecology and Evolutionary Biology

Defense Committee:
Scott A. Taylor, PhD, Department of Ecology and Evolutionary Biology
Barbara Demmig-Adams, PhD, Department of Ecology and Evolutionary Biology
Zoe Donaldson, PhD, Department of Molecular, Cellular, and Developmental Biology and Department of Psychology and Neuroscience

Abstract:

Cryptic divergence (i.e. evolution of reproductive isolation between populations that are morphologically similar) is an emerging target of population genetic research in the effort to better understand biodiversity in many ecosystems. Understanding the frequency of cryptic divergence and associated formation of new species, and the contexts in which these occur, helps establish where the resulting cryptic biodiversity (genetic diversity potentially not represented by visible phenotypic differences) may be most common. In birds, cryptic divergence and reproductive isolation in the absence of obvious morphological or behavioral differences appear common in the tropics but are not well-described in temperate regions. Here, we used reduced representation genome sequencing (ddRAD) to investigate possible cryptic divergence and reproductive isolation in a common North American songbird, the house wren (*Troglodytes aedon*). Evaluation of a combination of museum specimens and samples collected along an elevation transect in Boulder County, Colorado, provided evidence that the two North American subspecies of the house wren, distinguished by highly divergent mitochondrial genomes, are hybridizing along the Front Range in a previously unknown and undescribed cryptic contact zone. Both the eastern and western subspecies are present along the Rocky Mountain and Great Plain transition zone, with evidence for extensive hybridization across the genomes of Front Range house wrens. This work sets the stage for future investigations of whole genome differentiation and reproductive isolation in this newly identified North American avian hybrid zone.

Keywords: Hybridization, Hybrid Zone, *Troglodytes aedon*, Speciation

Introduction:

Population divergence and resulting formation of new species are complex, and often non-linear, processes influenced by many factors including physical isolation, genetic drift (random loss of genetic variation), natural selection (evolutionary process by which advantageous alleles persist in populations due to increased reproductive fitness), and gene flow (the transfer of genetic material between populations through interbreeding) (Harrison and Harrison 1993). Physical isolation and the resulting divergence, also known as allopatric speciation, is considered the most common process by which new species arise (Mayr 1942). When populations are physically isolated, genetic drift and natural selection can cause changes in allele frequencies that can eventually lead to morphological and behavioral differences (as shown for the white-breasted nuthatch and warbling vireo; Spellman and Klicka 2007; Lovell et al. 2021). In some cases, the events that caused the physical isolation can be reversed, causing the divergent populations to come back into secondary contact. Upon secondary contact, populations are sometimes similar enough to interbreed (Brelsford 2010; Toews et al. 2011), but have often accumulated enough differences that their admixed offspring suffer fitness costs (McQuillan et al. 2018). Differences that accumulate between isolated populations and play a role in reproductive isolation are known as barriers to reproduction. Such reproductive barriers can come in many forms, including prezygotic (pre-reproduction) barriers like differences in plumage, song, migration timing, reproductive physical mechanics, and habitat (Benites et al. 2010; Turbek et al. 2018; Xu and Shaw 2019; Jablonski et al. 2020; Cramer et al. 2021), as well as postzygotic (post-reproduction) barriers like reduced hybrid fitness and sterility

(Lanyon 1979; McQuillan et al. 2018). Ultimately, the degree of divergence that accumulates before secondary contact influences the degree of reproductive isolation or gene flow between populations and the outcomes of secondary contact.

Hybrid zones are often described as windows into the evolutionary process and provide a unique opportunity to understand generation and/or maintenance of reproductive barriers (Harrison and Harrison 1993). Individuals with a mixed genomic background in hybrid zones can be used to quantify the consequences of gene flow between divergent lineages. For example, the extent of admixture can (i) provide insight into whether hybridization results in reduced fitness and (ii) highlight genomic regions resistant to gene flow that may play a role in maintenance of species boundaries (McQuillan et al. 2018). Examining the underlying processes that lead to genetic divergence and, ultimately, reproductive isolation, is key to understanding the process of speciation and the generation of reproductive barriers. More broadly, hybrid zones provide unique opportunities to study the maintenance of species boundaries, and the evolutionary outcomes of genetic exchange between divergent lineages, and the study of North American avian hybrid zones has provided many important insights (Drovetski 2003; Cicero 2004; McQuillan et al. 2018).

In northern and southern temperate regions, the contemporary distributions of many species reflect expansion from their refugia (physically isolated areas of acceptable climate and habitat) during the last ice age in the Pleistocene period (Hewitt 2001). In North America, refugia typically existed beyond the southern edges of glaciers present along the contemporary border between the United States and Canada. For many species, isolation in refugia facilitated divergence via genetic drift and selection

(Avise and Walker 1998). After glacial retreat, numerous populations experienced secondary contact with close relatives, in some cases blending back into a single population, but in other cases forming regions of interbreeding and admixture, now visible as multigenerational hybrid zones. In North American birds, the outcomes of divergence and postglacial secondary contact vary (Klicka et al. 2011; Carling and Thomassen 2012; Lovell et al. 2021). Some species exhibit significant differences in plumage and song, as established for bullock's and baltimore orioles (Walsh et al. 2020), indigo and lazuli buntings (Sibley and Short 1959), and rose-breasted and black-headed grosbeaks (Mettler and Spellman 2009) that all have formed what appear to be stable regions of hybridization. In contrast, other systems exhibit few morphological or song differences, but exist as distinct populations that do not extensively interbreed, as is the case for warbling vireo (Lovell et al. 2021). For some species, like the warbling vireo, observational data from eBird, a citizen science-based observation repository, suggests that secondary contact and subsequent hybridization is primarily caused by movement of the western subspecies eastward from the mountains and onto the plains. Indeed, eastern forms of many species or subspecies pairs are rarely reported in the Rocky Mountains. This means that many of the Great Plains hybrid zones may be maintained due to western birds expanding into the plains and forming one-way tension hybrid zones (maintained by a balance of dispersal and selection), but this requires further investigation.

The avian species and subspecies pairs that populate the Rocky Mountain-Great Plains contact zone also are characterized by relatively strict adherence to specific elevational zones. The exact elevation that precludes the eastern subspecies is different

in each system, but the ranges of all of the hybridizing pairs do correlate with elevational changes. In some systems, including the hairy woodpecker, subspecies meet at the foothills of the Rockies (Klicka et al. 2011). The eastern subspecies in these types of systems can be found up to around 5,000 feet of elevation but not in the Rocky Mountains themselves. Other systems are much further removed from the abrupt transition of the plains to the Rocky Mountains, like the baltimore and bullock's oriole hybrid zone. This latter hybrid zone is centered around Western Kansas, with an elevational cline between 3,800 to 2,500 feet (Rising 1996). Both of these systems are distinctly tied to specific elevation zones and associated habitat types. The hairy woodpecker's hybrid zone is divided by the abrupt rise of the Rocky Mountains, while the oriole's hybrid zone is defined by the beginning of the geologic uplift of the High Plains.

In almost all pairs of avian species or subspecies, the partners are distinguishable by features that presumably act as visual (plumage) or vocal (song) reproductive barriers and reflect underlying genetic divergence (Kenyon et al. 2011; Seneviratne et al. 2012; Aguillon et al. 2021). Prior to genetic sequencing, hybridization was described as interbreeding between two populations with distinguishable and heritable morphological differences (Harrison and Harrison 1993). Modern genetic sequencing now allows exploration of whole-genome variation and characterization of species barriers and hybrid zones for genetically divergent populations that are morphologically indistinguishable but produce hybrids of intermediate genotypes at contact zones in North America (Slager et al. 2020). Patterns of morphological similarity with genetic divergence are also found in the neotropics, where morphologically similar

or identical taxa are often reproductively isolated and genetically distinct (cryptic) species (Pulido-Santacruz et al. 2018; Cronemberger et al. 2020). These cryptic species can have high levels of genetic diversity and rely on different niches, all while being visually and vocalling indistinguishable from one another. The cryptic nature of these systems stresses the importance of identifying divergent lineages and understanding the processes that lead to divergence. Characterizing the mechanisms that underlie such cryptic reproductive barriers could point to key differences in how morphologically similar species are able to diverge, and persist as cryptic species.

The house wren is a highly diverse species, with 32 identified subspecies and the largest breeding range of any songbird in the New World (Clements 2007). Across the continental United States, two distinct house wren mitochondrial clades are found, the Eastern and Western forms. These two forms exhibit 4% mitochondrial divergence, which is comparable to other species-level pairs (e.g., black-capped and carolina chickadees, 5% divergent Reudink 2007). The same unpublished mitochondrial data were used to create a phylogenetic tree, which presented results claiming that the eastern North America house wren subspecies was sister to a subspecies in Panama, and the western form was sister to a subspecies found in the west Mexican mountains. Broadly, these results indicate deep mitochondrial divergence within house wrens (Chavez et al. unpublished data), it is unclear whether this pattern is reflective of divergence in the nuclear genome, as other systems with significantly diverged mitochondrial genomes show limited levels of divergence in the nuclear genome (e.g., golden-winged and blue-winged warblers; Toews et al. 2016).

Our study sought (1) to determine whether both subspecies of North American

house wren occur along the Rocky Mountain-Great Plains transition zone (Front Range), and thereby support previous mitochondrial data and (2) advance our knowledge of the house wren system by quantifying and characterizing the contact zone and potential hybridization (admixture) therein. We obtained and sequenced both museum specimens and field-collected blood samples of house wrens from across the Front Range of Colorado, including individuals from high elevations in the Rocky Mountains and individuals from east of the foothills, out on the Great Plains.

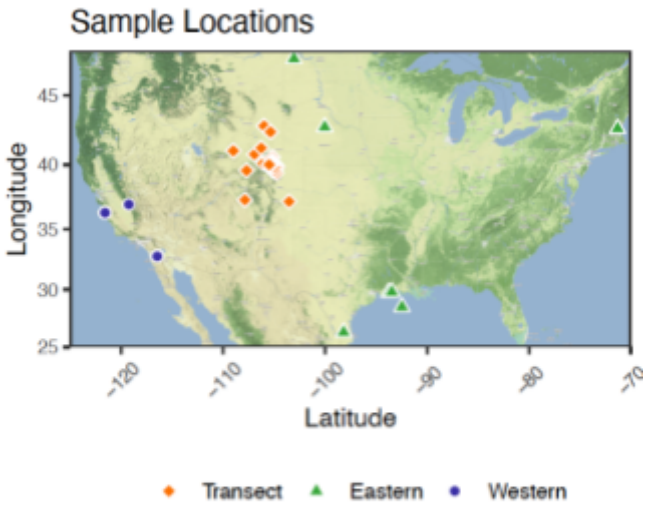

Figure 1. This is a map of the sampling locations of all 128 samples involved in this study. The western samples are shown in blue circle(n=9), Eastern samples in green triangles)n=9), and the contact samples are shown in orange diamonds(n=110)

Methods:

We sampled 128 house wrens, including allopatric western house wrens (n=9), allopatric eastern house wrens (n=9), and 110 birds from the contact zone along the Front Range of the Rocky Mountains in Colorado, USA (Figure 1). This included 39 birds (Table 1) from an elevational transect of nest boxes in Boulder, Colorado. This transect was created and maintained by the Boulder Chickadee Study, but every year, about 50 boxes are used by house wrens for breeding, allowing access to a dense elevational transect of wrens and their chicks. Boxes were monitored by the Boulder Chickadee Study field crew during the summer of 2021, who noted the stage of each nest, allowing us to sample fledglings in the nest. Fledglings were sampled and banded around an age

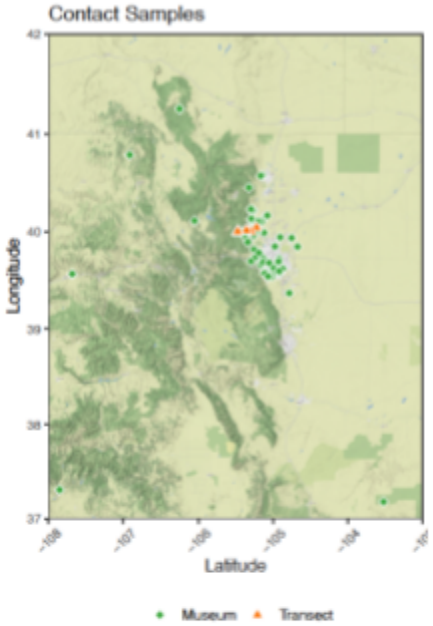

Figure 2. This is a map of the sampling distribution of all contact individuals sampled. Transect birds (n=39) are in orange triangles and museum specimens(n=71) are in green diamonds.

of 12-15 days. Approximately 40ul of blood was extracted from the brachial vein and was then stored at room temperature in lysis buffer until DNA extraction. The remaining 71 contact zone samples were obtained from museum collections (Table 2) and parental samples came from areas of allopatry (Table 3). Museum specimens used as contact zone samples were collected in Colorado and Wyoming (Figure 2). Genomic DNA was extracted from a combination of tissue and blood samples using a salt extraction protocol. Samples were first lysed using a homogenizing solution (0.4 M NaCl, 10 mM Tris–HCl pH 8.0, and 2 mM EDTA pH 8.0, Proteinase K), and a 20% SDS solution. We added 2 µl of glycoblue dye to aid in the identification of the DNA pellet, and precipitated DNA using a 6 M NaCl solution and 100% EtOH. DNA pellets were resuspended in 100 µl of TE buffer (10 mM Tris, 1 mM EDTA at pH 8–9). DNA concentrations were quantified using a Qubit fluorometer 3.0, using 2ul of DNA and 198ul of qubit solution.

After extraction, samples were sent to Admera Health for genomic library preparation. We used a double digest restriction site-associated reduced representation sequencing approach (ddRAD; Peterson et al. 2012). DNA samples were digested using the enzymes Sbfl/Mspl and were then ligated to an adaptor and a unique individual barcode. Digested fragments selected for sequencing were approximately 300 base pairs long. After digestion and ligation, selected samples were sequenced on

an Illumina NovaSeq with an S4 flow cell. Admera Health removed one individual due to low input, resulting in a return of 128 total samples (Parental n=18, Contact n=110).

Raw reads were trimmed using trimmomaticPE (Bolger et al. 2014) and were then aligned to a brown creeper (*Certhia americana)* chromosome level reference genome (NCBI ASM1869719v1) using the BWA mem algorithm (Li and Durbin 2010) with default settings. Variants were called with bcftools mpileup (Li et al. 2009), and were filtered to keep single nucleotide polymorphisms (SNPs) with a quality score greater than 80. SNPs were then filtered to include only SNPs with a mean depth above 0.5x and below 15x and a minimum minor allele frequency of 0.01 using VCFtools v 4.4 (Danecek et al. 2011). Multiple datasets were generated with the prior constraints, one allowing for 5% missing data (95p) and one allowing for 25% missing data (75p). Detailed instructions on the bioinformatic pipelines used are published on github (https://github.com/erikrfunk/whole_genome_bioinformatics).

Because many of our samples came from geographically proximate nest boxes, we calculated a relatedness statistic for each pair of individuals using an unadjusted Ajk statistic implemented in VCFtools. We retained one individual per pair, or group of individuals, with relatedness values greater than 0.25.

After generating the 95p and 75p datasets, we used VCFtools to create a subset of loci from the 75p dataset with $F_{ST} > 0.4$ that we could use to distinguish the eastern and western house wren subspecies. F_{ST} is a measure of population divergence that measures relative allele frequency differences and varies from 0 (no population difference) to 1 (populations fixed for different alleles). High F_{ST} is interpreted as high levels of divergence between subsets of a total population. Trimming for high values of

F_{ST} removes loci that are not informative and show no signal of divergence. The trimmed F_{ST} dataset contained the SNPs with the highest allele frequency differences between our allopatric parental populations allowing us to maximize the resolution of our ddRAD sequencing and provide a reasonable dataset for estimating genetic admixture between eastern and western house wrens. All analyses were conducted on this subset of SNPs with F_{ST} >0.4 unless otherwise noted.

To examine population genetic structure, we ran a principal component analysis (PCA) using the R program SNPRelate (Zheng et al. 2012). We used 88,826 SNPs from unrelated individuals, drawn from the 75p dataset, for this analysis.

We calculated hybrid indices (a measure of the degree of admixture within an individual's genome) for each individual from the contact zone using the R package gghybrid (v1.0.0, Bailey 2020). Samples from allopatric western and allopatric eastern localities were designated as non-admixed populations and were used to polarize genotypes. To ensure stability of the posterior, we ran three separate analyses with varying chain lengths including 10,000, 20,000, and 100,000 iterations, with corresponding burn-ins of 5,000, 10,000 and 50,000. All plots showed similar patterns, so we chose the plot with 100,000 iterations for increased confidence in the Bayesian Markov Chain Monte Carlo analysis, the sampling algorithm used by gghybrid. Parental populations were assigned a hybrid index of 0 (Western) or 1 (Eastern).

To test for relationships between the hybrid indices found using the gghybrid analysis and the elevation at which a specific individual was sampled, we ran linear regressions on our entire dataset, excluding the parental samples, and a subset of data only including our transect samples from Boulder County. Elevation was categorized as

the X (predicting) variable, and hybrid index was categorized as the Y (dependent) variable. This allowed us to test the hypothesis that elevation had a significant effect on hybrid index.

Results:

Following filtering, the 25% missing dataset (75p) contained 91,759 loci and the 5% missing (95p) contained 46,297 loci. Following additional filtering by high F_{ST} (>0.4), our final dataset contained 175 informative loci. Genome-wide variation among unrelated individuals summarized by a Principle Component Analysis (PCA) (generated with the 75p dataset, trimmed for relatedness, 88 individuals retained) showed two distinct parental populations separated along PC1 (1.39% of variation) corresponding to allopatric samples (Figure 3). Individuals from the contact zone fell between the parental populations, almost entirely along PC1.

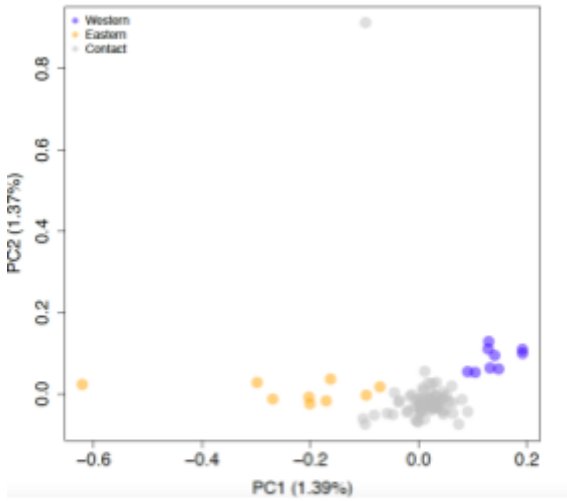

Figure 3. Principal Component Analysis of 88,826 SNPs. Western parental samples are in blue, eastern parental samples are in orange, and house wrens sampled throughout the Front Range of Colorado shown in gray.

We then ran the program gghybrid, using our F_{ST} >0.4 subset of 175 loci to produce hybrid indices for the population. Hybrid indices for the contact population calculated by gghybrid ranged from pure western (hybrid index = 0) to pure eastern (hybrid index = 1) (Figure 4). Furthermore, we found that individuals sampled from along

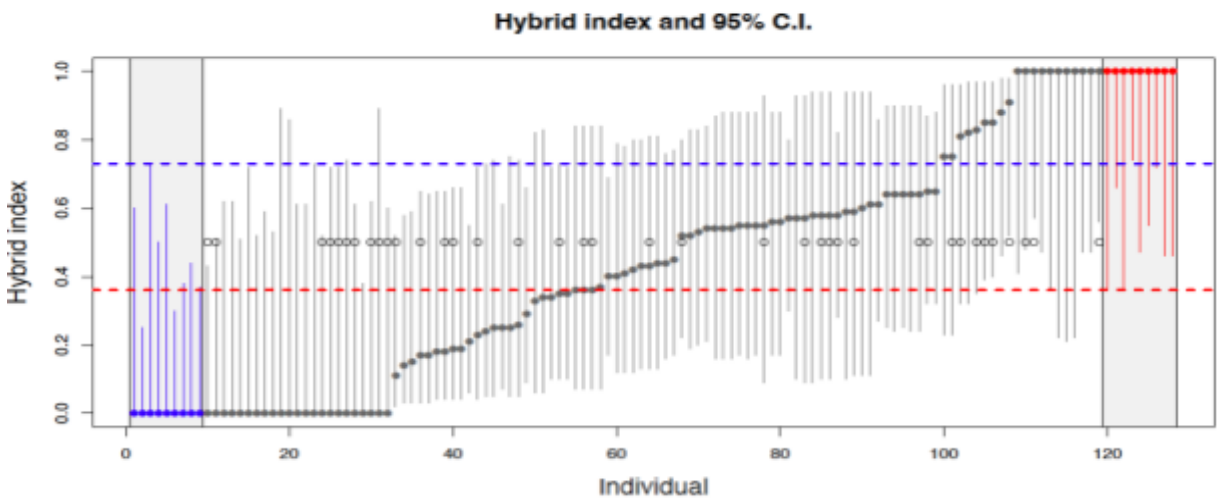

Figure 4. Gghybrid plot of all the individuals within the study (95p dataset). Western parental samples are shown in blue on the bottom left with a value of 0, eastern parental samples are shown in red with a value of 1 at the top right of the figure, individuals from the potential hybrid zone (i.e., the Front Range) are shown in gray, and individuals sampled from the transect in Boulder County are indicated with open circles.

the Boulder Chickadee Study transect (indicated with open circles in Figure 4) were assigned hybrid indices that included both pure western to pure eastern genomes, but also individuals that were assigned to a range of hybrid indices. This indicates that our transect (width=14 km) spans the width of the house wren contact zone.

The results of the regression analyses differed for the entire dataset and the subset of only Boulder Chickadee Study transect birds. There was no significant relationship between hybrid index and elevation ($F_{1,108}$= 0.471, p-value = 0.494). The regression examining the relationship between hybrid index and elevation and only samples collected along the Boulder Chickadee Study transect did, however, reveal a significant relationship between elevation and hybrid index ($F_{1,36}$= 8.022, p-value of 0.008).

Discussion:

The results of our study, i.e., identification of a hybrid zone between eastern and

western house wrens in North America and a sharp transition between subspecies and a range of hybrid indices across a 14-km nest box transect, support unpublished findings by Dr. Garth Spellman at the Denver Museum of Nature and Science that the two subspecies of house wren were about as divergent in their mitochondria as certain species-level pairs, like black-capped and carolina chickadees (Gill et al. 1993). Our nuclear data from reduced representation genome sequencing ddRAD suggest a lower, albeit significant, level of differentiation.

The short, 14km, nest box transect over which the transition from eastern to western house wrens takes place suggests that there is some degree of reproductive isolation, or reduced fitness in hybrids, that keeps the hybrid zone narrow as seen in other systems. For example, narrow hybrid zones are maintained by reduced hybrid fitness in the black-capped and carolina chickadee contact zone in Pennsylvania. In the latter system, hybrid individuals suffer from metabolic inefficiencies necessary for temperature regulation in the harsh Pennsylvania winter and learning and memory deficiencies that may interfere with their ability to survive the winter (Olson et al. 2010; McQuillan et al. 2018; Wagner et al. 2020). In our system, the broad range of hybrid indices we report indicates that many individuals are multigenerational wren hybrids that survive and reproduce. In this regard, the house wren hybrid zone is more similar to other Great Plains avian hybrid zones like spotted and eastern towhees (Sibley and West 1959) or bullock's and baltimore orioles (Allen 2002). Both of the latter systems feature multigenerational hybrid swarms, like the house wren, where the two parental populations come into contact, but differ from the house wren system in zone width. Generally, hybrid zones characterized by large multigenerational swarms much wider

than the house wren's 14-km transect, such as 112-482 km for the yellow-rumped warbler (Hubbard 1969) and ~400 km for the bullock's and baltimore oriole (Walsh et al. 2020).

Many of the Great Plains avian hybrid zones appear to be maintained due to selection against hybrids, but the mechanisms underlying this selection often remain unclear (Mettler and Spellman 2009). One potential explanation for the mechanism underlying divergence is a genetic basis for elevational adaptation. For example, based on tracked changes in precipitation over the last few decades, maintenance of the narrow hybrid zone for indigo x lazuli bunting hybrid zone appears to involve local adaptation to environmental variation (Carling and Thomassen 2012). Specifically, the transition from xeric (dry) to mesic (moist) conditions across the Great Plains also involves significant changes in elevation (Carling and Thomassen 2012). However, no significant association between elevation and hybrid index was found for the overall distribution of house wren subspecies and their hybrids across the hybrid zone when all data from the present study were included. In contrast, a significant association between hybrid index and elevation was obtained when only the samples collected across the nest box transect were included. The majority of museum samples that were included in the larger dataset came from lower elevations in the foothills and the greater Denver metro area. It is thus possible that the data from the latter metro area overwhelmed a signal resulting from a trend of western birds coming out of the mountains more than eastern birds going into the mountains. This latter trend would result in admixture of all levels (the whole span of hybrid indices) at a lower elevation, with changes only evident when samples collected in the mountains are included. Our

transect of elevation from Boulder, Colorado, to Nederland, Colorado (5318-8236 ft) excluded many of the lower elevation samples and thus removed the heavily mixed metro area. This interpretation, i.e., sampling along a transect is potentially more informative than inclusion of the entire dataset, but requires further testing via additional sampling.

Our initial assumption had been that the eastern form would be found exclusively on the plains because in other, similar species pairs, such as hairy woodpecker or baltimore and bullock's orioles, the western species or subspecies is found in both mountains and plains, but the eastern species or subspecies are not found in the mountains (Klicka et al. 2011; Walsh et al. 2020). The latter scenario suggests ecological or physiological features that allow the western forms to exist both at higher elevation and on the plains but prevent eastern species or subspecies from occurring at higher elevations. We originally expected the house wren system to exhibit similar patterns, but our data clearly show that the eastern form of the house wren is capable of dispersing into the mountains.

One possible explanation for this dispersion of the eastern house wren subspecies into the mountains could be habitat partitioning, the division of microhabitats between subspecies due to very specific habitat preferences. A study from the Andes mountains found that in areas of habitat-facilitated sympatry, two species of whitestart (genus *Myioborus*) were able to exist in sympatry due to habitat partitioning (Jablonski et al. 2020). The whitestarts used different types of habitats in areas of sympatry, allowing both to exist in the areas of elevational overlap. In our system, the eastern form of house wren is a riparian obligate, meaning that it requires riparian systems to

reproduce, not generally found in mountainous evergreen forests. However, we sampled hybrids at an elevation of 8,300 feet with a hybrid proportion of 0.58 (58% of individual's sampled loci assigned to the eastern genotype). It is possible that the distribution of the nest boxes along the transect used in the Boulder Chickadee Study artificially increased the samples of majority eastern genotype hybrids captured and sampled. Black-capped chickadees, one of the target taxa of the Boulder Chickadee Study, follow riparian corridors into the mountains. To maximize the number of black-capped chickadees sampled as elevation increases, boxes were distributed along these corridors. Since the eastern form of house wren also utilizes riparian systems, it is possible that these riparian corridors could thus act as paths for eastern birds to enter the mountains, in the same pattern as black-capped chickadees. Our sampling locations would thus be at a point of increased contact between the house wren taxa, increasing the odds of finding hybrids, and potentially inflating our estimates of rates of hybridization in the mountains.

The data generated in this project have created more questions pertaining to the house wren system. Observing high levels of differentiation in these two taxa leads to questions pertaining to their reproductive barriers. As mentioned above, it is possible that habitat partitioning is at least partially responsible for the genomic differentiation exhibited by house wrens. That said, we found that the high plains directly adjacent to the foothills of the Rocky Mountains contained individuals spanning the entire hybrid index. We also would like to further our understanding of the stability of the zone, looking for potential elevational movement. Many hybrid zones and species distributions are currently moving due to climate change, such as the black-capped and carolina

Chickadee zone (Taylor et al. 2014), and more locally, the bark beetle in the Colorado Rockies (Williams and Liebhold 2002; Hebertson and Jenkins 2008). Both of those systems are moving, north (chickadees) or higher in elevation (beetles), and could predict patterns in our house wren system. Climate-change mediated habitat disruption and movement has the potential to remove the potential barrier to reproduction that is habitat partitioning, resulting in increased introgression facilitated by increased access to mountainous habitats by eastern taxa.

Another potential barrier to reproduction could be entirely beyond the habitat or genetic mechanisms of the house wren, focused entirely on the temporal patterns exhibited by these migrant songbirds. Many species of North American migratory birds have two populations, one that migrates over land and one that migrates over the ocean. For example, the black-throated blue warbler migrates and winters in two different populations, and this difference in pattern and wintering ground has resulted in significant genetic divergence (Davis et al. 2006). It has been postulated that temporal differences, which can arise due to different barriers across differing migration routes, like the timing of migration, leads to temporal isolation and subsequent divergence due to drift and assortative mating (Turbek et al. 2018). It has also been noted that in some species, migratory patterns have been reduced or stopped, resulting in rapid speciation due to the temporal isolation pertaining to breeding time (Gómez-Bahamón et al. 2020). Those factors could be influencing house wren divergence, and would explain the lack of divergence across the plumage morphology and vocalizations.

It is important to note that the resolution of our ddRAD nuclear data is not ideal

for this investigation. Using this low-resolution approach was beneficial from a monetary perspective, but the confidence intervals provided by gghybrid were very large due to the low number of loci used (i.e., 175 SNPs). Future work on the house wren hybrid zone will increase genomic resolution by using a whole genome sequencing approach. We would also like to drastically increase the size and scope of the study transect as to more comprehensively characterize the hybrid zone. Expanding boxes into areas beyond riparian zones will allow us to characterize habitat that should be inhabited by pure western individuals and will allow us to further our understanding of this hybrid zone system.

In summary, the present study establishes the Front Range of the Rocky Mountains as the approximate western boundary for the Western house wren, and the eastern boundary for the Eastern house wren. Extensive, multi-generational hybridization between the two lineages, with differing patterns of hybridization on the plains and in the mountains, is evident. These patterns are similar to other Great Plains avian hybrid zones but lead to further questions about the nature of the reproductive boundaries that maintain what appears to be a relatively narrow hybrid zone. More research is needed to further our understanding of the cryptic processes underlying the house wren subspecies hybrid zone.

Acknowledgements:

I would like to thank Dr. Scott Taylor, Erik Funk, and the entire Taylor Lab for their support and guidance throughout this entire experience, I would not have been able to do it without them. I would also like to thank Dr. Garth Spellman and the Denver

Museum of Nature and Science for his assistance in the formation of this study and his specimen contributions. I could not have been able to collect all of the samples used in this study without the help of Angela Theodosopoulos and the entire Boulder Chickadee Study Field Crew, as well as all of the landowners who graciously allow us to utilize their property for nest box placement. I would also like to thank the University of Washington's Burke Museum and Louisiana State University's Museum of Natural History for their generous specimen loan contributions. Finally, I would like to thank the Denver Field Ornithologists, Colorado Field Ornithologists, and the Biological Sciences Initiative for their generous funding of this project.

References:

Aguillon SM, Walsh J, Lovette IJ. 2021. Extensive hybridization reveals multiple coloration genes underlying a complex plumage phenotype. Proc R Soc B Biol Sci. 288(1943):20201805. doi:10.1098/rspb.2020.1805.

Allen ES. 2002. Long-term hybridization and the maintenance of species identity in orioles (*Icterus*). Indiana University.

Avise JC, Walker DE. 1998. Pleistocene phylogeographic effects on avian populations and the speciation process. Proc R Soc Lond B Biol Sci. 265(1395):457–463. doi:10.1098/rspb.1998.0317.

Bailey RI. 2020. gghybrid: R package for evolutionary analysis of hybrids and hybrid zones. Zenodo. [accessed 2022 Jan 28]. https://zenodo.org/record/3676499.

Benites P, Eaton MD, Lijtmaer DA, Lougheed SC, Tubaro PL. 2010. Analysis from avian visual perspective reveals plumage colour differences among females of capuchino seedeaters (*Sporophila*). J Avian Biol. 41(6):597–602. doi:10.1111/j.1600-048X.2010.05205.x.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 30(15):2114–2120. doi:10.1093/bioinformatics/btu170.

Brelsford A. 2010. Hybridization and speciation in the yellow-rumped warbler complex. University of British Columbia. [accessed 2021 Oct 21]. https://open.library.ubc.ca/soa/cIRcle/collections/ubctheses/24/items/1.0071052.

Carling MD, Thomassen HA. 2012. The role of environmental heterogeneity in maintaining reproductive isolation between hybridizing passerina (*Aves*:

Cardinalidae) buntings. Int J Ecol. 2012:e295463. doi:10.1155/2012/295463.

Cicero C. 2004. Barriers to sympatry between avian sibling species (*paridae*: *Baeolophus*) in local secondary contact. Evolution. 58(7):1573–1587. doi:10.1111/j.0014-3820.2004.tb01737.x.

Cramer ERA, Grønstøl G, Maxwell L, Kovach AI, Lifjeld JT. 2021. Sperm length divergence as a potential prezygotic barrier in a passerine hybrid zone. Ecol Evol. 11(14):9489–9497. doi:10.1002/ece3.7768.

Cronemberger ÁA, Aleixo A, Mikkelsen EK, Weir JT. 2020. Postzygotic isolation drives genomic speciation between highly cryptic *Hypocnemis* antbirds from Amazonia. Evolution. 74(11):2512–2525. doi:10.1111/evo.14103.

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, et al. 2011. The variant call format and VCFtools. Bioinformatics. 27(15):2156–2158. doi:10.1093/bioinformatics/btr330.

Davis LA, Roalson EH, Cornell KL, Mcclanahan KD, Webster MS. 2006. Genetic divergence and migration patterns in a North American passerine bird: implications for evolution and conservation. Mol Ecol. 15(8):2141–2152. doi:10.1111/j.1365-294X.2006.02914.x.

Drovetski SV. 2003. Plio-Pleistocene climatic oscillations, holarctic biogeography and speciation in an avian subfamily. J Biogeogr. 30(8):1173–1181. doi:10.1046/j.1365-2699.2003.00920.x.

Gill FB, Mostrom AM, Mack AL. 1993. Speciation in North American chickadees: i. patterns of mtDNA genetic divergence. Evolution. 47(1):195–212. doi:10.1111/j.1558-5646.1993.tb01210.x.

Gómez-Bahamón V, Márquez R, Jahn AE, Miyaki CY, Tuero DT, Laverde-R O, Restrepo S, Cadena CD. 2020. Speciation associated with shifts in migratory behavior in an avian radiation. Curr Biol. 30(7):1312-1321.e6. doi:10.1016/j.cub.2020.01.064.

Harrison Richard Gerald, Harrison Richard Granville. 1993. Hybrid zones and the evolutionary process. Oxford University Press.

Hebertson EG, Jenkins MJ. 2008. Climate factors associated with historic spruce beetle (*Coleoptera*: *Curculionidae*) outbreaks in Utah and Colorado. Environ Entomol. 37(2):281–292. doi:10.1093/ee/37.2.281.

Hewitt GM. 2001. Speciation, hybrid zones and phylogeography — or seeing genes in space and time. Mol Ecol. 10(3):537–549. doi:10.1046/j.1365-294x.2001.01202.x.

Hubbard JP. 1969. The relationships and evolution of the *Dendroica coronata* complex. The Auk. 86(3):393–432. doi:10.2307/4083405.

Jablonski PG, Borowiec M, Nowakowski JJ, Stawarczyk T. 2020. Ecological niche partitioning in a fragmented landscape between two highly specialized avian flush-pursuit foragers in the Andean zone of sympatry. Sci Rep. 10(1):22024. doi:10.1038/s41598-020-78804-2.

Clements, James F. 2007. The Clements Checklist of Birds of the World. 6th ed. Cornell University Press.

Kenyon HL, Toews DPL, Irwin DE. 2011. Can song discriminate between macgillivray's and mourning warblers in a narrow hybrid zone? The Condor. 113(3):655–663. doi:10.1525/cond.2011.100182.

Klicka J, Spellman GM, Winker K, Chua V, Smith BT. 2011. A phylogeographic and

population genetic analysis of a widespread, sedentary North American bird: the hairy woodpecker (*Picoides villosus*). The Auk. 128(2):346–362. doi:10.1525/auk.2011.10264.

Lanyon WE. 1979. Hybrid sterility in meadowlarks. Nature. 279(5713):557–558. doi:10.1038/279557a0.

Li H, Durbin R. 2010. Fast and accurate long-read alignment with Burrows–Wheeler transform. Bioinformatics. 26(5):589–595. doi:10.1093/bioinformatics/btp698.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 25(16):2078–2079. doi:10.1093/bioinformatics/btp352.

Lovell SF, Lein MR, Rogers SM. 2021. Cryptic speciation in the warbling vireo (*Vireo gilvus*). The Auk.:ukaa071. doi:10.1093/ornithology/ukaa071.

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

McQuillan MA, Roth TC, Huynh AV, Rice AM. 2018. Hybrid chickadees are deficient in learning and memory. Evolution. 72(5):1155–1164.

Mettler RD, Spellman GM. 2009. A hybrid zone revisited: molecular and morphological analysis of the maintenance, movement, and evolution of a Great Plains avian (*Cardinalidae*: *Pheucticus*) hybrid zone. Mol Ecol. 18(15):3256–3267. doi:10.1111/j.1365-294X.2009.04217.x.

Olson JR, Cooper SJ, Swanson DL, Braun MJ, Williams JB. 2010. The relationship of metabolic performance and distribution in black-capped and carolina chickadees.

Physiol Biochem Zool PBZ. 83(2):263–275. doi:10.1086/648395.

Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE. 2012. Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. PLOS ONE. 7(5):e37135. doi:10.1371/journal.pone.0037135.

Pulido-Santacruz P, Aleixo A, Weir JT. 2018. Morphologically cryptic Amazonian bird species pairs exhibit strong postzygotic reproductive isolation. Proc R Soc B Biol Sci. 285(1874):20172081. doi:10.1098/rspb.2017.2081.

Reudink M. 2007. Structure and dynamics of the hybrid zone between black-capped chickadee (*Poecile atricapillus*) and carolina chickadee (*P. carolinensis*) in southeastern Pennsylvania | Ornithology | Oxford Academic. Ornithology. [accessed 2021 Oct 6]. https://academic.oup.com/auk/article/124/2/463/5562685.

Rising JD. 1996. The stability of the oriole hybrid zone in western kansas. The Condor. 98(3):658–663. doi:10.2307/1369584.

Seneviratne SS, Toews DPL, Brelsford A, Irwin DE. 2012. Concordance of genetic and phenotypic characters across a sapsucker hybrid zone. J Avian Biol. 43(2):119–130. doi:10.1111/j.1600-048X.2012.05516.x.

Sibley CG, Short LL. 1959. Hybridization in the buntings (*Passerina*) of the Great Plains. The Auk. 76(4):443–463. doi:10.2307/4082312.

Sibley CG, West DA. 1959. Hybridization in the rufous-sided towhees of the Great Plains. The Auk. 76(3):326–338. doi:10.2307/4081810.

Slager DL, Epperly KL, Ha RR, Rohwer S, Wood C, Van Hemert C, Klicka J. 2020. Cryptic and extensive hybridization between ancient lineages of american crows.

Mol Ecol. 29(5):956–969. doi:10.1111/mec.15377.

Spellman GM, Klicka J. 2007. Phylogeography of the white-breasted nuthatch (*Sitta carolinensis*): diversification in North American pine and oak woodlands. Mol Ecol. 16(8):1729–1740. doi:10.1111/j.1365-294X.2007.03237.x.

Taylor SA, White TA, Hochachka WM, Ferretti V, Curry RL, Lovette I. 2014. Climate-mediated movement of an avian hybrid zone. Curr Biol. 24(6):671–676. doi:10.1016/j.cub.2014.01.069.

Toews DPL, Brelsford A, Irwin DE. 2011. Hybridization between townsend's *Dendroica townsendi* and black-throated green warblers *D. virens* in an avian suture zone. J Avian Biol. 42(5):434–446. doi:10.1111/j.1600-048X.2011.05360.x.

Toews DPL, Taylor SA, Vallender R, Brelsford A, Butcher BG, Messer PW, Lovette IJ. 2016. Plumage genes and little else distinguish the genomes of hybridizing warblers. Curr Biol. 26(17):2313–2318. doi:10.1016/j.cub.2016.06.034.

Turbek SP, Scordato ESC, Safran RJ. 2018. The role of seasonal migration in population divergence and reproductive isolation. Trends Ecol Evol. 33(3):164–175. doi:10.1016/j.tree.2017.11.008.

Wagner DN, Curry RL, Chen N, Lovette IJ, Taylor SA. 2020. Genomic regions underlying metabolic and neuronal signaling pathways are temporally consistent in a moving avian hybrid zone. Evolution. 74(7):1498–1513. doi:10.1111/evo.13970.

Walsh J, Billerman SM, Rohwer VG, Butcher BG, Lovette IJ. 2020. Genomic and plumage variation across the controversial baltimore and bullock's oriole hybrid zone. The Auk. 137(4). doi:10.1093/auk/ukaa044. [accessed 2021 Oct 19].

https://doi.org/10.1093/auk/ukaa044.

Williams DW, Liebhold AM. 2002. Climate change and the outbreak ranges of two North American bark beetles. Agric For Entomol. 4(2):87–99. doi:10.1046/j.1461-9563.2002.00124.x.

Xu M, Shaw KL. 2019. The genetics of mating song evolution underlying rapid speciation: linking quantitative variation to candidate genes for behavioral isolation. Genetics. 211(3):1089–1104. doi:10.1534/genetics.118.301706.

Zheng X, Levine D, Shen J, Gogarten SM, Laurie C, Weir BS. 2012. A high-performance computing toolset for relatedness and principal component analysis of SNP data. Bioinformatics. 28(24):3326–3328. doi:10.1093/bioinformatics/bts606.

Table 1: Boulder Chickadee Study Transect Individuals

Admera ID	Location	Hybrid Index	Elevation
94	40.04562, -105.2197	0.83	5209
95	40.04562, -105.2197	0	5209
96	40.04562, -105.2197	1	5209
97	40.04562, -105.2197	0.64	5209
98	40.04562, -105.2197	0.58	5209
99	40.04562, -105.2197	0.91	5209
100	40.04562, -105.2197	0.17	5209
75	40.00394, -105.29019	0.44	5676
76	40.00394, -105.29019	0.25	5676
101	40.01684, -105.35469	0.52	6478
102	40.01684, -105.35469	1	6478
103	40.01684, -105.35469	0.81	6478
104	40.01684, -105.35469	0.85	6478
105	40.01684, -105.35469	0.35	6478
106	40.01684, -105.35469	0.65	6478
107	40.01684, -105.35469	1	6478
108	40.00378, -105.46189	0.19	8081
109	40.00378, -105.46189	0.12	8081
110	40.00378, -105.46189	0	8081
111	40.00378, -105.46189	0	8081
112	40.00378, -105.46189	0.23	8081
89	39.994, -105.47851	0.36	8280
90	39.994, -105.47851	0	8280

91	39.99395, -105.47851	0.58	8280
92	39.99395, -105.47851	0	8280
93	39.99395, -105.47851	0.19	8280
77	40.00631, -105.47223	0	8348
78	40.00631, -105.47223	0	8348
79	40.00631, -105.47223	0	8348
80	40.00631, -105.47223	0.56	8348
81	40.00631, -105.47223	0	8348
82	40.00631, -105.47223	0.57	8348
83	40.00631, -105.47223	0.86	8348
84	40.00631, -105.47209	0.36	8348
85	40.00631, -105.47209	0	8348
86	40.00631, -105.47209	0.75	8348
87	40.00631, -105.47209	0.61	8348
88	40.00631, -105.47209	0.58	8348

Table 2. Museum Samples

Admera ID	Location	Hybrid Index	Elevation
1	37.18161, -103.52113	0.42	5936
2	39.9392, -104.75116	0.14	5106
3	39.3707, -104.78212	1	6363
4	39.97488, -105.2628	0.57	5634
5	39.89636, -105.33605	0	7849
6	39.72839, -104.92423	0.18	5357
7	40.20502, -105.27509	0.17	5717

8	40.11394, -105.27669	0.55	5509
9	40.015366, -105.324276	0	5750
10	40.11557, -106.04827	0.29	8098
11	39.84879, -104.67285	0.64	5440
12	40.054369, -105.248769	0.4	5345
13	39.57695, -105.05918	0.35	5452
14	40.00441, -105.40442	0.54	6892
15	40.78471, -106.91559	1	7965
16	39.96815, -105.25534	1	5715
17	41.03815, -108.98814	0.88	9352
18	41.03815, -108.98814	1	9352
19	41.03815, -108.98814	1	9352
20	41.03815, -108.98814	0	9352
21	39.54306, -105.06482	0.36	5457
22	39.706892, -105.133833	0.19	5783
23	41.25045, -106.24808	0.15	9894
24	39.57534, -105.11216	0.52	5762
25	39.62102, -104.98258	0.41	5448
26	39.58128, -104.90756	0	5665
27	39.66225, -105.1598	0.25	5495
28	40.5774, -105.13336	0	5139
29	39.54291, -105.01334	0.56	5643
30	39.56839, -107.68477	0.54	6000
31	39.60083, -104.98912	0.34	5506
32	40.10807, -105.154	0.22	5058
33	39.56839, -107.68477	0.58	5627

34	39.9933, -105.12009	0.54	5358
35	39.73127, -105.13345	0.44	5639
36	40.12854, -105.29169	0.36	5878
37	40.16935, -105.07615	0.6	5040
38	39.54306, -105.06482	0.45	5457
39	39.70243, -105.29278	0.44	8224
40	39.62214, -104.86748	0	5656
41	40.22459, -105.26501	0.81	5410
42	39.81876, -105.2512	0.33	6171
43	39.63792, -104.9988	0.66	5371
44	40.04906, -105.24624	0.35	5347
45	39.68948, -104.94383	0	5413
46	40.11406, -105.27686	0	5508
47	40.57309, -105.16337	1	5433
48	40.01535, -105.28167	1	5351
49	37.30966, -107.85424	0	6554
50	37.30966, -107.85424	0.64	6554
51	39.74512, -105.22927	0	6009
52	39.78743, -105.18822	0.43	5665
53	39.68882, -105.13267	0	5832
54	40.01533, -105.32408	0.41	5752
55	40.45298, -105.32197	0.75	7092
56	40.13575, -105.28902	1	5723
57	40.01465, -105.20665	0.6	5246
58	40.03626, -105.28252	0.64	5469
59	39.69778, -104.92512	0.54	5357

60	39.96511, -105.3829	0.55	7947
61	40.10777, -105.20312	0.56	5155
62	39.68229, -105.05317	0.26	5440
63	39.59814, -104.9181	0.54	5579
120	42.85551, -105.97297	0.25	5003
121	42.85551, -105.97297	0	5003
122	42.85551, -105.97297	0.24	5003
123	42.42862, -105.34472	0.57	6456
124	42.42862, -105.34472	0.64	6456
125	42.42862, -105.34472	0.6	6456
126	39.94631, -104.90426	0.53	5221
127	39.85232, -104.97427	0	5221
128	40.23447, -105.29346	0.61	5514

Table 3. Parental Samples

Admera ID	Location	Hybrid Index	Population
64	32.78503, -116.45091	0	Western
65	36.305000, -121.568333	0	Western
66	36.305000, -121.568333	0	Western
67	36.305000, -121.568333	0	Western
68	32.851333, -116.420000	0	Western
69	36.958333, -119.231667	0	Western
70	36.305000, -121.568333	0	Western
72	36.305000, -121.568333	0	Western
74	32.78503, -116.45091	0	Western

71	42.6688, -71.302	1	Eastern
73	47.5032, -103.043	1	Eastern
113	28.5125, -92.4511	1	Eastern
114	29.75383, -93.75821	1	Eastern
115	29.79919, -93.44537	1	Eastern
116	26.30642, -98.17371	1	Eastern
117	26.30642, -98.17371	1	Eastern
118	42.78524, -100.02767	1	Eastern
119	42.78341, -100.02749	1	Eastern