ABSTRACT

BEHAVIORAL AND GENETIC DIVERGENCE AMONG WILD AND DOMESTICATED POPULATIONS OF THE ZEBRA FINCH (*TAENIOPYGIA GUTTATA*)

by

Allison L. Lansverk

July 2017

Director of Dissertation: Christopher N. Balakrishnan, Department of Biology

Interdisciplinary Doctoral Program in Biological Sciences

ABSTRACT

Zebra finches (*Taeniopygia guttata*) have been the subject of extensive neurological and behavioral research having served as the dominant model for vocal learning over half a century. Learned vocal communication, or vocal learning, is a trait that is shared by humans and songbirds but is rare or less well developed in other animals. Unlike innate communication, learned vocalizations are acquired early on by juveniles listening and copying what they hear from adults. Little, however, has been done to characterize the intraspecific variation in song behavior in the zebra finch model system. Other systems, such as the lab mouse, *Mus musculus*, have begun to take advantage of inbred and natural populations to assess genetic variation and to link genotype and behavior. The opportunity exists to do the same in the zebra finch. The first step to better able study song learning in a genetics context is to define trait variation within and among populations. The majority of research conducted on these birds relies on domesticated populations of *Taeniopygia guttata castanotis* (*T. g. castanotis*), but wild populations are also

available for study, as is a second subspecies, *T. g. guttata*. With the sequencing of the zebra finch genome a decade ago, zebra finches have risen in importance in the field of population genomics so there is an opportunity to investigate the genetic variation in this system as well. I compared patterns genetic and song variation among these populations to examine how these features have diverged during the early stages of domestication as well as during divergence in allopatry. When comparing the wild and domesticated populations, I find that overall levels of genetic differentiation are low (F_{ST} = ~0.02); I also find evidence of selection acting on portions of the genome. Genetic drift also appears to have played a role in shaping patterns of genetic variation. While genetic drift has led to reduced diversity and a loss of rare alleles in domestic populations, it has also done so in the island subspecies, *T. g. guttata*: I found further support for a dramatic bottleneck in the island subspecies as the two subspecies have diverged, as there is an overall reduction in diversity. Among the most highly diverged regions of the genome are two genes associated with color. I have identified fixed differences in two well-known pigmentation genes, SLC45A2 and CDKN2A that may contribute to plumage color differences between subspecies. In addition to genetic divergence, I also characterized divergence in song behavior among populations. I find that the island subspecies shows less variation in song among individuals than the mainland birds. Though the island subspecies, *T. g. guttata*, shows a reduction in variation in song among individuals possibly due to the bottleneck during speciation, the domestication process has actually led to increased variability in song structure in domesticated birds. It is possible that domesticated birds have been freed from the constraints on song structure imposed by mate choice and the need for accurate species recognition. Finally, in order to differentiate between genetic or cultural controls of this difference in variation, I cross-fostered both subspecies to the Bengalese finch, *Lonchura striata domestica*, to test for

differences in song copying behavior. I cannot reject the null hypothesis that zebra finch subspecies copy tutor songs equally well, but it does appear that the high variability in song structure in *T. g. castanotis* remains following controlled tutoring. Overall, I have begun to characterize the intraspecific behavioral and genetic variation in zebra finches, which has the potential to further our ability to study gene-environment influences on behavior, particularly with regards to the genetic contributions to song copying ability.

BEHAVIORAL AND GENETIC DIVERGENCE AMONG WILD AND DOMESTICATED POPULATIONS OF THE ZEBRA FINCH (*TAENIOPYGIA GUTTATA*)

A Dissertation

Presented To

The Faculty of the Interdisciplinary Doctoral Program in Biological Sciences,
The Brody School of Medicine

In Association with the Department of Biology,
Thomas Harriot College of Arts and Sciences

East Carolina University

In Partial Fulfillment of the Requirement for the Degree

Doctor of Philosophy

Interdisciplinary Doctoral Program in Biological Sciences

By

Allison L. Lansverk

July 2017

© Copyright 2017
Allison L. Lansverk

BEHAVIORAL AND GENETIC DIVERGENCE AMONG WILD AND DOMESTICATED POPULATIONS OF THE ZEBRA FINCH (*TAENIOPYGIA GUTTATA*)

By

Allison L. Lansverk

APPROVED BY:

DIRECTOR OF DISSERTATION: _____
Christopher N. Balakrishnan, Ph.D.

COMMITTEE MEMBER: _____
Susan B. McRae, Ph.D.

COMMITTEE MEMBER: _____
Jeffery S. McKinnon, Ph.D.

COMMITTEE MEMBER: _____
Carol Goodwillie, Ph.D.

COMMITTEE MEMBER: _____
Kenneth M. Soderstrom, Ph.D.

COMMITTEE MEMBER: _____
Michael S. Brewer Ph.D.

CHAIR OF THE DEPARTMENT OF BIOLOGY: _____
Jeffery S. McKinnon, Ph.D.

DEAN OF THE GRADUATE SCHOOL: _____
Paul J. Gemperline, Ph.D.

DEDICATION

To my parents for their unlimited love and support. Without them I never would have realized this dream.

To my greatest ally David for his constant encouragement and for believing in me always.

ACKNOWLEDGEMENTS

"No man is an island entire of itself; every man is a piece of the continent, a part of the main"

– John Donne

No Ph.D. is an island either. This body of work would not have been possible without the tireless effort and support of many people. First I would like to say the biggest thank you to Chris Balakrishnan for being such a fantastic advisor and mentor. Advisors can make or break a dissertation, and you surely helped make mine. Thank you for never giving up on me even when I wanted to give up on myself. Through a challenging field season and subsequent project change, thank you for being supportive and helping me brainstorm new directions for my Ph.D. Your love of science is certainly inspiring, and no matter where my path takes me I hope that passion will be something I will always have for science as well. And thanks to Jeff McKinnon for putting me in contact with Chris in the first place all those years ago.

To the rest of the Balalab, Dan, John, Dustin, Matt, and Robert, thank you for being great colleagues and even better friends. Through the troubleshooting code, to taking blood samples from zebra finches, to enjoying our tradition of Indian buffet lunches, you all make the day to day on campus more fun.

Thank you to my committee members, Sue McRae, Jeff McKinnon, Carol Goodwillie, Ken Soderstrom, and Michael Brewer for the many suggestions and words of encouragement over the years. A special thanks to Michael for his patience and tireless support helping me troubleshoot all sorts of bioinformatics issues over the years. I would probably still be fighting the first steps of ANGSD if not for you. And thanks to Ken for showing me how to do the song

analyses portions of my project. I know I would never have made it through those analyses without your help.

There is a long list of collaborators that contributed their data or time to this work. Thanks to Sarah London for the recordings from her birds. Thanks to Simon Griffith for his collaboration and his access to his zebra finch populations. Thanks to Sonal Singhal for sequence data and for being there to answer questions. Thanks to Thorfinn Korneliussen, Matteo Fumagalli, Alan Bergland, Zev Kronenberg, Dean Ousby, Matteo Fumagalli, and Ofer Tchernichovski for help with various software and scripts. And thanks to Jean-Bernard Dogmo and Justin Schuetz and Eric Fishel for being there for my ill-fated pilot field season of my original dissertation project.

A big thanks to everyone in IDPBS, especially Terry West and Ed Stellwag. Also to the wonderful staff over the years, especially Barb Beltran, Joyce Beatty, Tammy McKinney, Julie Marik, and Jen Jacobs. Also thanks to the staff who keep the building clean and everything running, who also manage to be so kind and supportive of the students. I also must extend my sincere thanks to the vets and techs over at Brody, especially Karen Oppelt, Dorcas O'Rourke, Matt Verzwyvelt, and Julie Briley. Thank you for taking care of the birds and always being there whenever there were any issues.

My friends and family helped me keep my spirits up even through stressful times of my Ph.D. Thanks to Lauren McCarthy, Suelen Tullio, Debbie Lichti, Nikki Davis Armstrong, Kelly Mahoney, as well as many more friends within the biology family and outside of it. Thank you to all of the fantastic biology teachers that I've had over the years, spanning middle school, high school, and college: you all helped me fall in love with science in the first place, and I wouldn't have made it to this point without you. A huge thank you to my parents, Kay and Marvin, and

my sister Hannah for all their love and constant support over the years. I would probably never even have started my Ph.D. let alone ever finished it without you. And thank you dad, for reading what is surely every draft of every paper I've ever written. Finally, a special thank you goes to David. Thank you for believing in me when I often didn't believe in myself and encouraging me every step of the way.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1 GENETIC AND BEHAVIORAL CONSEQUENCES OF RECENT DOMESTICATION IN THE ZEBRA FINCH (*TAENIOPYGIA GUTTATA*)

Abstract

Zebra finches (*Taeniopygia guttata*) have been the subject of extensive behavioral and neurological research. The majority of research done on these birds relies on domesticated populations, yet little is known about patterns of genomic and behavioral divergence between domesticated and wild populations. Here I investigate variation in the genomes and song production of domesticated and wild zebra finches. Although I find that overall levels of genetic differentiation are low (F_{ST} = ~0.02), I also find evidence of selection acting on portions of the genome. Genetic drift also appears to have played a role in shaping patterns of genetic variation and song production. While drift has led to reduced diversity and a loss of rare alleles in domestic populations, it has also led to increased variability in song structure in domesticated birds. Domesticated birds have likely been freed from the constraints on song structure imposed by mate choice and the need for accurate species recognition. These differences between wild and domesticated birds may have important implications for future studies of vocal behavior in zebra finches.

Introduction

The selective and demographic influences of animal domestication result in dramatic phenotypic and genetic changes in the species we choose to domesticate (Rubin et al. 2010, Shapiro et al. 2013, Vickrey et al. 2015, but see Poh et al. *2014*). These changes have been well studied in many of the classic model systems, like the domesticated chicken, *Gallus gallus domesticus* (Rubin et al. 2010) and the house mouse, *Mus musculus* (Estep et al. 1975). Domestication is often accompanied by specific changes such as a reproductive period that is uncoupled from the natural environment, as well as changes in body size and coloration (Trut et al. 2009). As it is likely these physiological traits are affected during domestication due to drift or selection, these same forces can act on behavioral traits (Price 1999). The domesticated chicken, *Gallus gallus domesticus*, which has been selected for traits that result in high meat or egg production, shows reduction in behaviors that were energetically costly (therefore reserving energy for production) relative to their wild counterparts (Schütz and Jensen 2001). Domestic mice show more frequent social encounters than their wild counterparts (Smith et al. 1994), and reduced food/water neophobia (Kronenberger and Medioni 1985), and are less effective at avoiding predation (Kardong 1993). Since domestication affects genomes and behaviors, it is important to understand these impacts in widely used model systems.

The zebra finch *Taeniopygia guttata* is a model system for studying behavior. As a representative of the Oscine Passerines, or songbirds, this species has been the subject of extensive neurobiological and behavioral research (Jin and Clayton 1997; Dave and Margoliash 2000; Olveczky et al. 2005; Mooney 2009; London and Clayton 2008; Thompson et al. 2011; Vallentin et al. 2016). For over half of a century, the *T. g. castanotis* subspecies of zebra finch has been the dominant model for vocal learning (Forstmeier et al. 2009; Slater et al. 1988) and an

important model for understanding the origins of the complexity of the human language (Marler 1970; ten Cate 2014). The majority of research conducted on these birds relies on domesticated populations, yet little is known about patterns of genomic divergence between domesticated and wild populations (but see Forstmeier et al. 2007) or patterns of variation in song behavior between these two populations. Unlike many domesticated species, captive zebra finch populations are often kept as outbred colonies, with relatively free mating (Forstmeier et al. 2007). Also, the domestication of the zebra finch, *T. g. castanotis* subspecies, has been fairly recent (~150 years), and wild populations are still available for study, enabling analyses of rapid adaptation (Rogers 1979; Clayton 1989, Zann 1996), so an understanding of genetic variation in zebra finch populations is critical to further advancing the zebra finch as a model system for studying gene-behavior relationships.

Domestication is known to influence aspects of social behavior such as reduced aggression and fear response, increased social cognitive abilities for interacting with humans, and reduced wild-type behavior towards humans (Frank and Frank 1982; Schütz et al. 2001; Hare et al. 2002; Trut et al. 2009). Domestication impacts song learning and song production as well (Suzuki et al. 2004). Previous studies using different populations already suggest that there may be important differences between captive study populations, even those descended from the same wild-caught founders. For example, in some song learning studies, the number of elements that could not be assigned to a tutor has been found to be as low as 12% (Mann and Slater 1995) and as high as 50% (Jones et al. 1996; Houx and ten Cate 1999; Holveck et al. 2008). There has been some documentation of laboratory populations exhibiting discrete vocal traditions from one another (Sturdy et al. 1999) whereas the opposite has also been found (Lachlan et al. 2016). Additionally, the drawbacks of solely focusing on domesticated birds for behavioral research

have lately been garnering attention because recent studies have found that wild and domesticated zebra finches differ in an array of behaviors such as mate choice, hatching synchrony, nest visit synchrony, and parental care (Rustein et al. 2007; Mainwaring et al. 2010; Gilby et al. 2013; Mariette and Griffith 2012; Gilby et al. 2011).

In this chapter, I compare variation in the genomes and vocal behavior of between wild and domestic Australian zebra finches (*T. g. castanotis*). I find that overall genetic divergence is relatively low between wild and domesticated population, confirming some earlier work (Forstmeier et al. 2007). I also find evidence of selection acting on portions of the genome, likely a result of selective breeding by humans and adaptation to captivity. I find that selection and genetic drift have impacted patterns of both genomic and behavioral divergence in the domesticated zebra finch.

Methods

Genomic Divergence

Sample collection and sequencing. In order to derive a complete picture of genetic variation, a total of 39 unrelated individuals (20 domesticated and 19 wild zebra finches) were collected and sequenced. To avoid bias from sampling from one potentially inbred population, domesticated Zebra Finches were sampled from four laboratory populations in the United States that conduct active research on these birds: University of Illinois, University of Chicago, USGS National Wildlife Health Center, and East Carolina University. Permissions were obtained for the samples and they were collected with prior approval from the relevant ethical committees. For each zebra finch, genomic DNA was extracted from liver tissue using a DNeasy Blood and Tissue Kit (Qiagen kit cat no. 69504). Paired-end sequencing libraries with an insert size of 300 base pairs

were constructed following Illumina protocol for sequencing on the Illumina HiSeq 2000 platform. This platform was used to re-sequence paired end full genomes at medium coverage (8x) for 18 of the 20 domesticated zebra finches.

Nineteen wild zebra finch genomes and two domesticated zebra finches were sequenced using the Illumina HiSeq platform at medium coverage (20x) by collaborators Singhal et al. (2015). The 19 wild birds (10 females and 9 males) were collected from a studied population from Fowlers Gap in New South Wales, Australia (Singhal et al. 2015). Of the two domesticated birds that were sequenced by Singhal and colleagues (2015), one was a male and one was a female. From the 19 wild birds, DNA was extracted using a DNeasy Blood and Tissue Kit (Qiagen kit cat no. 69504). A species pooled library was diluted to a 10nM and sequenced across 15 lanes using 100 paired end reads with an Illumina HiSeq 2000. This was done in 2012 at the Oxford Genomics Centre at the Wellcome Trust Centre for Human Genetics in Oxford, UK. For the two domesticated birds, DNA was extracted from liver tissue using a DNeasy Blood and Tissue Kit (Qiagen kit cat no. 69504). A pooled library was made and sequenced across six lanes on 2-lane flow cells of an Illumina HiSeq using 100 base pairs paired-end reads. This was done in 2013 at the University of Chicago Genomics Core, Chicago IL USA (Singhal et al. 2015).

Sequence quality checking. For quality filtering, all the raw reads were trimmed using the program Trim Galore at default settings (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/).

Read mapping. The trimmed reads were aligned to the *Taeniopygia guttata* reference genome (accession number: WUGSC 3.2.4/taeGut1, Warren et al. 2010) using BWA-MEM (Li and

Durbin 2009) with default parameters. The 19 wild samples and 2 domesticated samples from Singhal et al. (2015) were then downsampled using SAMtools (version 1.2, using htslib 1.2.1) so the number of reads would be comparable to the remaining 18 domesticated samples. Next duplicates were removed using MarkDuplicates in Picard Tools v1.115. Following that RealignerTargetCreator and IndelRealigner in GATK v3.1-1 were used to conduct local realignment around indels. The next step was to fix the errors in the mate pairs using FixMateInformation in Picard Tools (http://broadinstitute.github.io/picard/). For each individual, this generated a realigned Binary sequence Alignment/Map (BAM) file.

Sex-linked scaffolds. Using SAMtools I removed the Z chromosomes (one of the zebra finch sex chromosomes) for the analysis that included all of the birds (male and female). I also separately analyzed only the male samples (the homogametic sex in birds) with the Z chromosome included.

SNP and genotype calling. I did single nucleotide polymorphism (SNP) calling using ANGSD (Korneliussen et al. 2014). The genotype likelihood method employed in ANGSD to estimate allele frequencies was expected to be especially useful because I cannot unambiguously assign genotypes to individuals with medium coverage genomes.

ANGSD uses a two-step procedure to estimate the site frequency spectrum (SFS), which is a distribution of allele frequencies for the set of SNPs. First sample allele-frequency likelihood files (.saf) were generated with the –doSaf 1 option with the ancestral state as that of the long-tailed finch, *Poephila acuticauda*, my out-group. Second, the .saf files were optimized using the

realSFS program to estimate the SFS. Genotypes were called from the genotype likelihood data. I also made VCFs using the tag –doVCF in ANGSD.

Genome-wide patterns of heterozygosity. Watterson estimate (y_w), Tajima's D statistic, and population-differentiation statistic (F_{ST}) were calculated using a sliding window approach of 100 base pair window sliding in 1000 base pair steps as well as a site-by-site estimation. For the analysis that included the Z chromosome, a method to calculate F_{ST} (-whichFst) that took into account small sample sizes was used.

Screening for outlier F_{ST} sites. I used the SegmentFst method in vcflib to find continuous regions with high F_{ST} values (vcflib – https://github.com/vcflib/vcflib). The tool SegmentF_{ST} scans the output of wcF_{ST} that I obtained from ANGSD and measures the number of F_{ST} values that are above the threshold I set of 0.8. I used the tool permuteSmooth to look at the empirical significance of the segmented F_{ST}. This tool takes the segmented F_{ST} data and shuffles them across the genome to regions with the same number of F_{ST} values where the empirical probability is calculated as the number of random trials where the average F_{ST} was higher than the observed segment. I then used the Integrative Genome Viewer (Robinson et al. 2011; Thorvaldsdóttir et al. 2013) to look at these sites to see if they were derived in the domesticated birds and look for anything noteworthy.

Screening for selective sweeps. Secondly, I used OmegaPlus to identify selective sweeps. OmegaPlus utilizes linkage disequilibrium (LD), whether a given haplotype is overrepresented in the population, to detect selective sweeps. The presence of strong LD might indicate that there

has been a recent selective sweep, and therefore can be used to identify sites under recent selection. I ran OmegaPlus with a -minwin of 200 and a -maxwin of 200000 and a -grid of 50000.

Investigation into potential enhancer region. I looked into three SNPs with extremely high F_{ST} located 300 kb downstream of PRL (prolactin). I ran two *in silico* promoter searches (ElemeNT – http://lifefaculty.biu.ac.il/gershon-tamar/index.php/element-description/element and NNPP – http://www.fruitfly.org/seq_tools/promoter.html). I BLATed against the pin-tailed whydah transcriptome, *Vidua macroura*, (ovaries + brain) to see if it was transcribed. I also checked whether this SNP cluster was different in domesticated *T. g. castanotis* compared to wild *T. g. castanotis*, *T. g. guttata*, pin-tailed whydah (*Vidua macroura*), long-tailed finch (*Poephila acuticauda*), and village indigobird (*Vidua chalybeata*).

Analyses of Song Structure

In order to characterize the variation in the patterns of song behavior between different populations, I analyzed songs from two domesticated and one wild zebra finch populations. Domesticated *T. g. castanotis* were recorded at East Carolina University and University of Chicago. Wild *T. g. castanotis* were recorded at Macquarie University in Sydney, Australia. Additionally, five captive male Bengalese finches, *Lonchura striata domestica*, from ECU were recorded for an outgroup comparison.

Study Populations. The East Carolina University population of domesticated *T. g. castanotis* birds was established in 2012. Originally, the domesticated *T. g. castanotis* colony was founded

from five pairs of birds derived from another captive research colony at ECU. All birds were housed in a large flight aviary at the Brody School of Medicine where they could freely feed on a commercial finch dry-seed mix, with fresh water provided daily and finely crushed warmed hard-boiled eggs and greens given weekly. The birds were monitored on a daily basis. Housing and experimental protocols were approved by the East Carolina University Animal Ethics Committee. Ten *T. g. castanotis* and five *L. s. domestica* were used for this study.

Collaborator Sarah London recorded eight domesticated *T. g. castanotis* zebra finches at the University of Chicago using the same methods I employed, and provided me with the recordings. The seven wild *T. g. castanotis* are from populations housed by my collaborator Simon Griffith in his lab at Macquarie University in Sydney, Australia. These birds were taken from the wild in 2007 from Sturt National Park in far northwest New South Wales and were allowed to breed in captivity in a couple of large aviaries. In 2010 a further set of birds were taken from Fowlers Gap (far-west NSW) and added to the population in captivity. In total about 100 adults were taken from Sturt with about 40 adults taken from Fowlers Gap. They have been isolated from domesticated birds, and have had the opportunity to breed about every 12 months (Gilby et al. 2013). In the summer of 2016, Chris Balakrishnan recorded them using the same method utilized for the birds at ECU.

Song Recordings. Pairs of birds, with one male and one female each, were placed in the sound chamber and left overnight to be recorded using Sound Analysis Pro (SAP) software (Tchernichovski et al. 2000, http://ofer.sci.ccny.cuny.edu/html/sound _analysis.html). The study of song learning has relied heavily on the development of software to analyze recordings and produce sonograms to look at similarities and differences between songs (Mooney 2009).

Tchernichovski and colleagues were able to investigate how a young bird's song differed note-to-note from the tutor's song, pioneering this type of study (Tchernichovski et al. 2001). SAP allows for activity-triggered recordings, making it easier to select out recording with vocalizations. The chamber was opened once a day to check food and water and the pair was left in the chamber for an average of three days or until the males produced at least 100 song recordings. I used a male and female instead which means that I captured both directed (to the female) and undirected song (not specifically directed to the female). Undirected song is slightly more diverse within-individual than directed, but Woolley and Doupe (2008) suggest that the within-individual differences are very subtle. Therefore, following previous work that has grouped both type of song together (Lachlan et al. 2016) I do not expect this to influence my overall results. Once the recordings were collected, I manually sorted through the files to select out only those that contained songs and discarded other vocalizations or noises from movement.

Sound Analysis Pro was also used for quantitative comparisons of song structure. As a measure of song similarity, I used the Kullback-Leibler divergence in pairwise comparisons among individuals (Wu et al. 2008). A higher K-L distance indicates more spread in the data and therefore songs that are more dissimilar among the individuals being compared. In cases where there are high levels of stereotypy, this can indicate that there is a high success in learning (Deregnaucourt et al. 2005). To accomplish this, the Feature Batch function in SAP was first used to parse the motifs into syllables by setting certain segmentation values, which are unique to each individual. To do this, I opened around 20 songs in SAP per individual and adjusted the segmentation setting for one or two features until each syllable was more effectively separated out. The features that can be used to segment are amplitude, pitch, mean frequency, goodness of pitch, FM, AM, Wiener entropy, continuity (t), and continuity (f). In most cases amplitude,

sometimes with a secondary feature of mean frequency or continuity, was used to segment the song recordings. Once the settings are selected for segmentation, the program automatically segments all the song recordings in a folder.

All K-L divergence estimates include syllable duration as one of the variables, but were estimated for a suite of different secondary song parameters. The batch analysis from SAP creates syllable tables which give information for 14 parameters: duration, mean amplitude, mean pitch, mean FM, mean AM^2, mean entropy, mean pitch goodness, mean frequency, variance in pitch, variance in FM, variance in entropy, variance in pitch goodness, variance in mean frequency, and variance in AM. One of the wild *T. g. castanotis* individuals did not have the necessary number of recordings in order for the K-L comparison to run, so its 44 song recordings were split in half using RAVEN (Bioacoustics Research Program 2011) to bring the number of recordings closer to 100 (note: the analysis was run with and without this individual and results did not differ if it was included or not).

Statistical Analyses. All statistical analyses were performed using the software R (R Core Team 2014). Statistical analyses were performed on approximately 100 songs per individuals. After calculating the K-L distance for the 13 song features, I ran a Shapiro-Wilk Normality Test on the data. Since most of the data were not normally distributed, I used a Mann-Whitney Wilcoxon test to determine if the K-L distances for each feature were significantly different for the subspecies comparison. I compared the ECU *T. g. castanotis* and University of Chicago *T. g. castanotis* to the wild *T. g. castanotis*. I also compared the ECU *T. g. castanotis* to the ECU *L. striata domestica*.

Results

Genomic Divergence

Genomic polymorphism & Divergence in wild and domesticated zebra finches. I detected 99,018,650 polymorphic sites between wild and domesticated zebra finches. As might be expected, the two populations differ significantly in diversity; diversity in domesticated zebra finches is reduced ~30% from population wild zebra finches ($p<0.001$). Average *theta* for wild zebra finches was 14.9 whereas average *theta* for domesticated zebra finches was 10.44. This difference is consistent across the genome with ~98% of 1000bp windows showing lower diversity in domesticated versus wild. The Watterson Estimator, *theta*, is estimated by the number of polymorphic sites. It is the product of the neutral mutation rate and the effective population size to show overall population mutation rate. Thus, higher value for *theta* is associated with greater diversity. Average Tajima's D also differed significantly between populations (Tajima's $D_{domesticated}$ = -0.75, Tajima's D_{wild} = -1.51, $p<0.01$). The value for Tajima's D becomes more negative when there is an excess of rare alleles. Along with the overall loss in diversity, there has been a loss of rare alleles, seen in the shift in the SFS (Figure 1.1). The wild birds have a higher proportion of alleles that are only represented once (Figure 1.1).

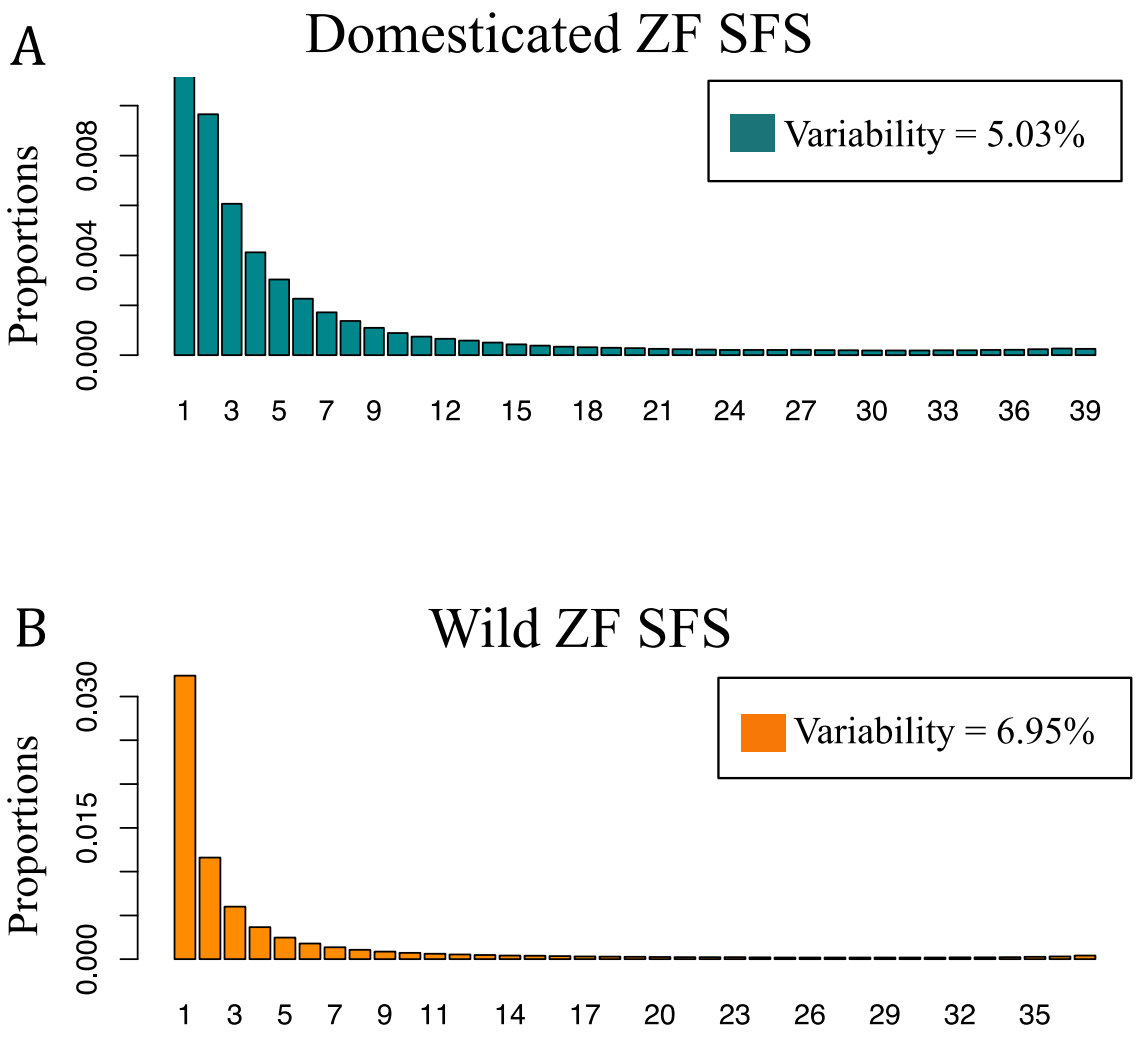

Figure 1.1. Site frequency spectra for wild and domesticated zebra finches.

The wild populations show a relative excess of rare alleles. The number of bins on the X-axis is equal to 1 minus the number of alleles in each population (39 for the 20 domesticated birds, and 37 for the 19 wild birds). The proportion represents the proportion alleles that are present one, or two, or three and more times.

I find that these populations show a low level of overall genetic differentiation (F_{ST} = 0.019), but in the sliding windows approach, F_{ST} ranged up to ~0.2, and in the site-by-site approach to ~0.9, indicating regions of high differentiation.

Selection during domestication. Despite overall low F_{ST}, it is possible that artificial selection has led to divergence in specific areas of the genome (Figure 1.2). I used multiple approaches to examine selection. I identified 12 outlier regions of high F_{ST} using the program SegmentFst from vcflib (Figure 1.2-1.3, Table 1.1).

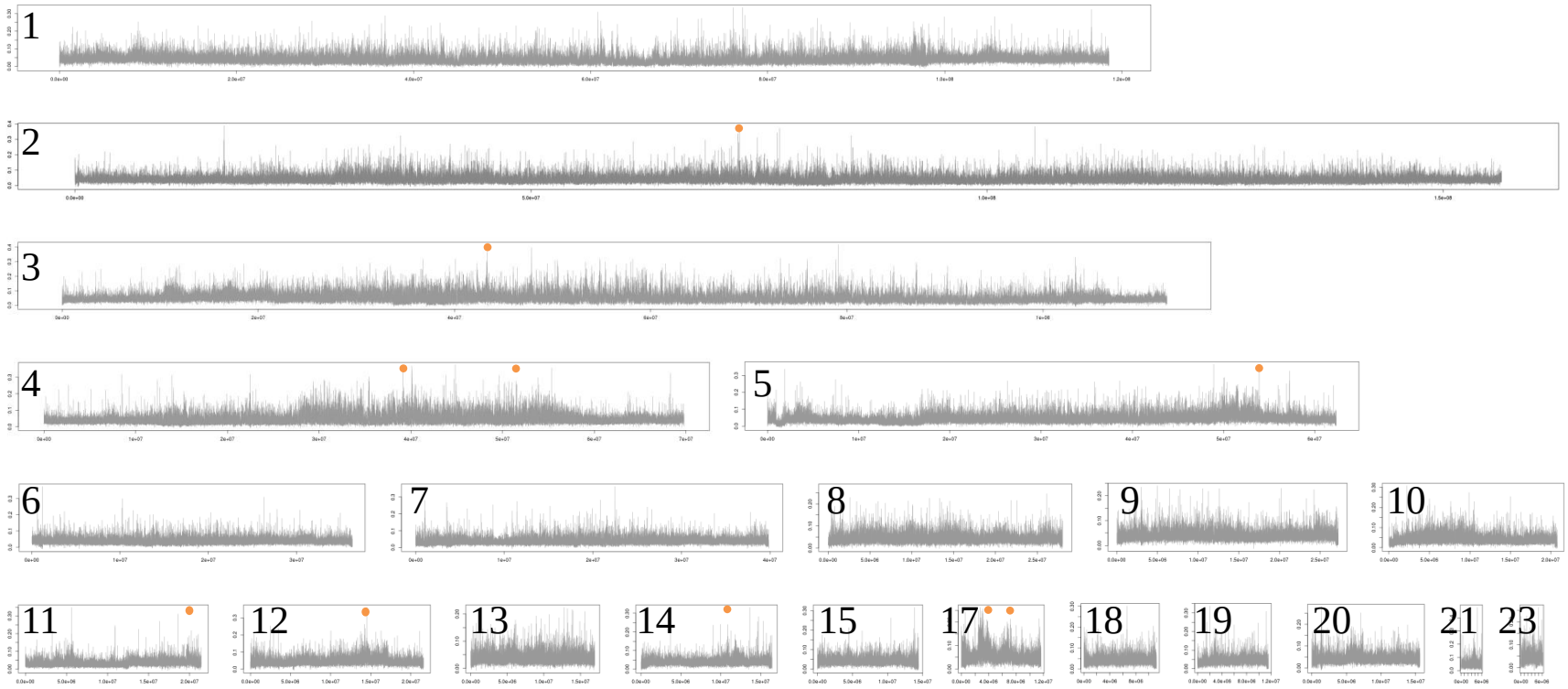

Figure 1.2. Windowed F_{ST} by chromosome.

Windowed F_{ST} (1000 base pair windows by 100 base pair steps) for each chromosome. Outlier SNPs identified with SegmentFst are marked in orange (these regions are highlighted in Figure 1.3).

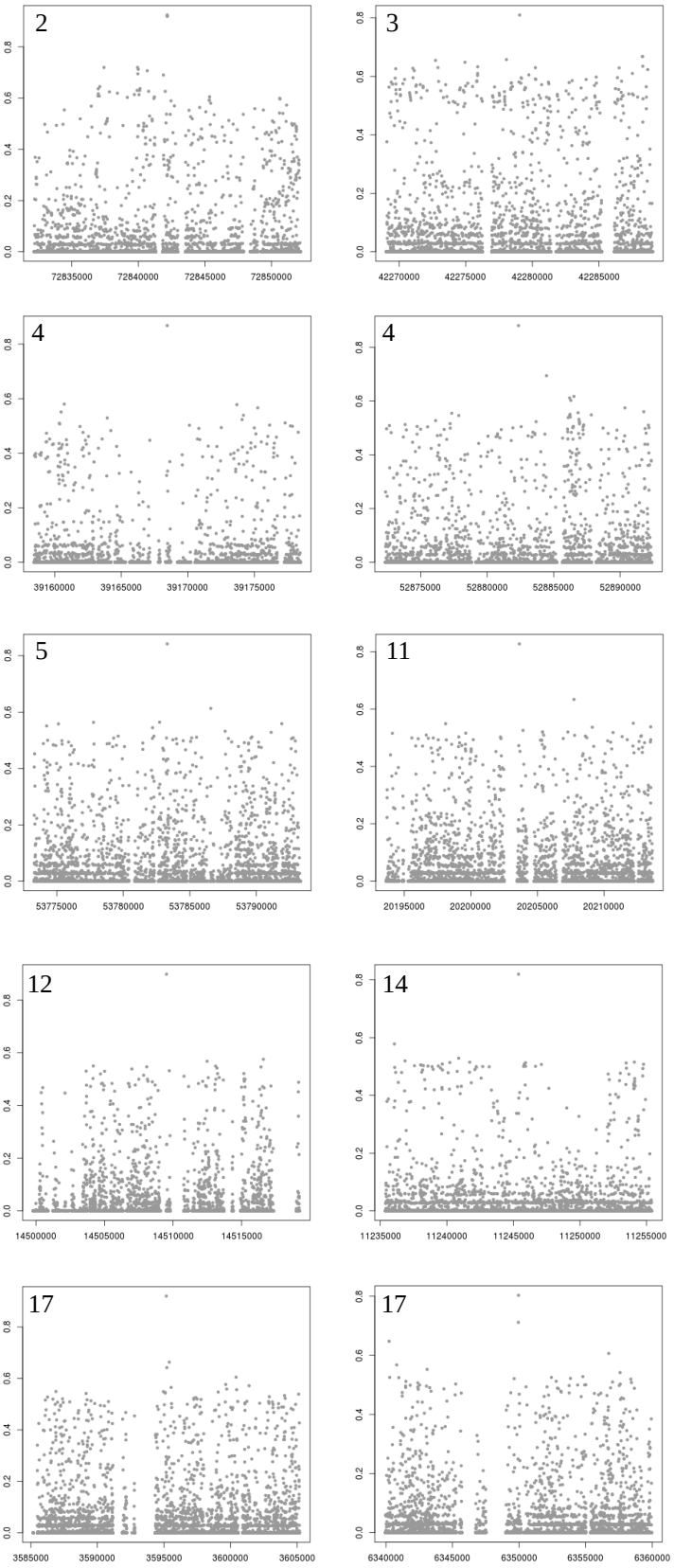
2
0.0
0.2
0.4
0.6
0.8
72835000
72840000
72845000
72850000
3
0.0
0.2
0.4
0.6
0.8
42270000
42275000
42280000
42285000
4
0.0
0.2
0.4
0.6
0.8
39160000
39165000
39170000
39175000
4
0.0
0.2
0.4
0.6
0.8
52875000
52880000
52885000
52890000
5
0.0
0.2
0.4
0.6
0.8
53775000
53780000
53785000
53790000
11
0.0
0.2
0.4
0.6
0.8
20195000
20200000
20205000
20210000
12
0.0
0.2
0.4
0.6
0.8
14500000
14505000
14510000
14515000
14
0.0
0.2
0.4
0.6
0.8
11235000
11240000
11245000
11250000
11255000
17
0.0
0.2
0.4
0.6
0.8
3585000
3590000
3595000
3600000
3605000
17
0.0
0.2
0.4
0.6
0.8
6340000
6345000
6350000
6355000
6360000

Figure 1.3. Outlier F_{ST} sites by chromosome.

20,000 base pair window of site-by-site F_{ST} around the outlier SNPs identified by SegmentFst on chromosomes 2, 3, 4, 5, 11, 12, 14, and17.

Table 1.1. Position and information about the outlier F_{ST} sites.

Outlier F_{ST} SNPs with information about where they are located (intergenic or genic), which genes they are in or near, and whether there is expression found in blood/brain (BB), whole brain (WB), or spleen (SP).

SNP Position	F_{ST}	Location	In or near gene	Uniquely derived in domesticated?	Expression
chr2_72842171	0.92	Intergenic region	CDKAL1-PRL	Yes	BB, WB
chr2_72842172	0.92	Intergenic region	CDKAL1-PRL	Yes	BB, WB
chr2_72842173	0.92	Intergenic region	CDKAL1-PRL	Yes	BB, WB
chr3_42279042	0.81	Intron variant	CAPN9	No	None
chr4_39168453	0.87	Intergenic region	SORBS2-TLR3	Yes	None
chr4_52882354	0.88	Intron variant	CCDC149	Yes	BB
chr5_53783303	0.84	Intergenic region	GPR132-JAG2	Yes	SP
chr11_20203641	0.83	Intergenic region	CDH5-ENSTGUG0000000984	Yes	None
chr12_14509519	0.90	Intron variant	PSMD6	Yes	BB, SP
chr14_11245379	0.82	Upstream gene variant	SMURF1	Yes	None
chr17_3595178	0.92	Downstream gene variant	ENSTGUG00000003327	Yes	None
chr17_6349955	0.80	Intron variant	FNBP1	No	SP

These regions of high F_{ST} are primarily intergenic. GO groups included categories such as regulation of multicellular organism growth, hormone activity, and signaling.

I used the Integrative Genome Viewer to look at all of the 12 sites from Table 1.1. Of the 12 sites, 10 were uniquely derived in the domesticated zebra finch compared to the wild zebra finch and the out-group, the long-tailed finch (Table 1.1). The region around the SNP at chr14:11245379, which falls a little under 5,000 base pairs upstream from the gene SMURF1, has much higher coverage in the wild zebra finches (average of 66 reads) than it does in the domesticated zebra finches (average of 5 reads). This might suggest that some sort of rearrangement has occurred. SMURF1 acts as a negative regulator of bone morphogenetic protein signaling pathway and has been identified as a candidate gene in a domestication studies on Yakutian horses (Librado et al. 2015). It has been put forth as a candidate gene for globularization and language readiness (Benítez-Burraco et al. 2016) as well as a region where a selective sweep occurred comparing Northern and Southern Europeans (Chen et al. 2010).

Three of the SNPs with the highest F_{ST} are all downstream of the gene that regulates prolactin (PRL). PRL is known as the parental hormone. In breeding birds an increase in prolactin is associated with periods of parental care (Angelier and Chastel 2009). There is preliminary evidence to suggest that those three nucleotides on chromosome two (Table 1.1) are a promoter. So, if this region is indeed associated with PRL, this could be a change in regulation, e.g. tissue-specific variant. Alternatively, this could be an enhancer or other transcription factor with unknown regulatory function (i.e. this region could affect distant genes and can even be found downstream). When I BLATed this region against the pin-tailed whydah transcriptome, *Vidua macroura*, (ovaries + brain) a region nearby <100 base pairs was transcribed.

Furthermore, those three nucleotides were derived in the domesticated *T. g. castanotis* (Table 1.1). This SNP cluster could be important, potentially to PRL transcription, but I was unable to confirm that it is indeed functionally important.

Secondly, I used OmegaPlus to identify selective sweeps. OmegaPlus utilizes linkage disequilibriums (LD), whether a given haplotype is overrepresented in the population, to detect selective sweeps. The presence of strong LD might indicate that there has been a recent selective sweep, and therefore can be used to identify sites under recent selection. After an FDR correction was done on the omega p-values, no sites were significant.

My findings highlight that genetic divergence is potentially due to both selection and genetic drift and point to putative functional differences between wild and domesticated populations of zebra finches.

Patterns of variation in song behavior

Genetic drift can influence patterns of behavioral variation as it does genetic variation. It is possible that bottlenecks might reduce variation, as they often do with genetic variation (Lacy 1997; Tsutsui et al. 2000). Thus, I wanted to investigate whether there was any difference in the variability of song production between wild and domesticated birds. For each bird that I recorded I captured between 3000 and 9000 syllables (N = 25 birds). I found that domesticated *T. g. castanotis* show increased variation between individuals compared with wild *T. g. castanotis* (Figure 1.4, row 1-3). Within the parameter plots, each cluster of dots represents a note that is repeatedly sung. Therefore, the similarity in the overall pattern of clustering between individuals visually represents how similar they are to each other. For example, for the domesticated *T. g.*

castanotis, bird 1A shows a similar overall pattern to bird 1B but is less similar to 1D (Figure 1.4, row 1).

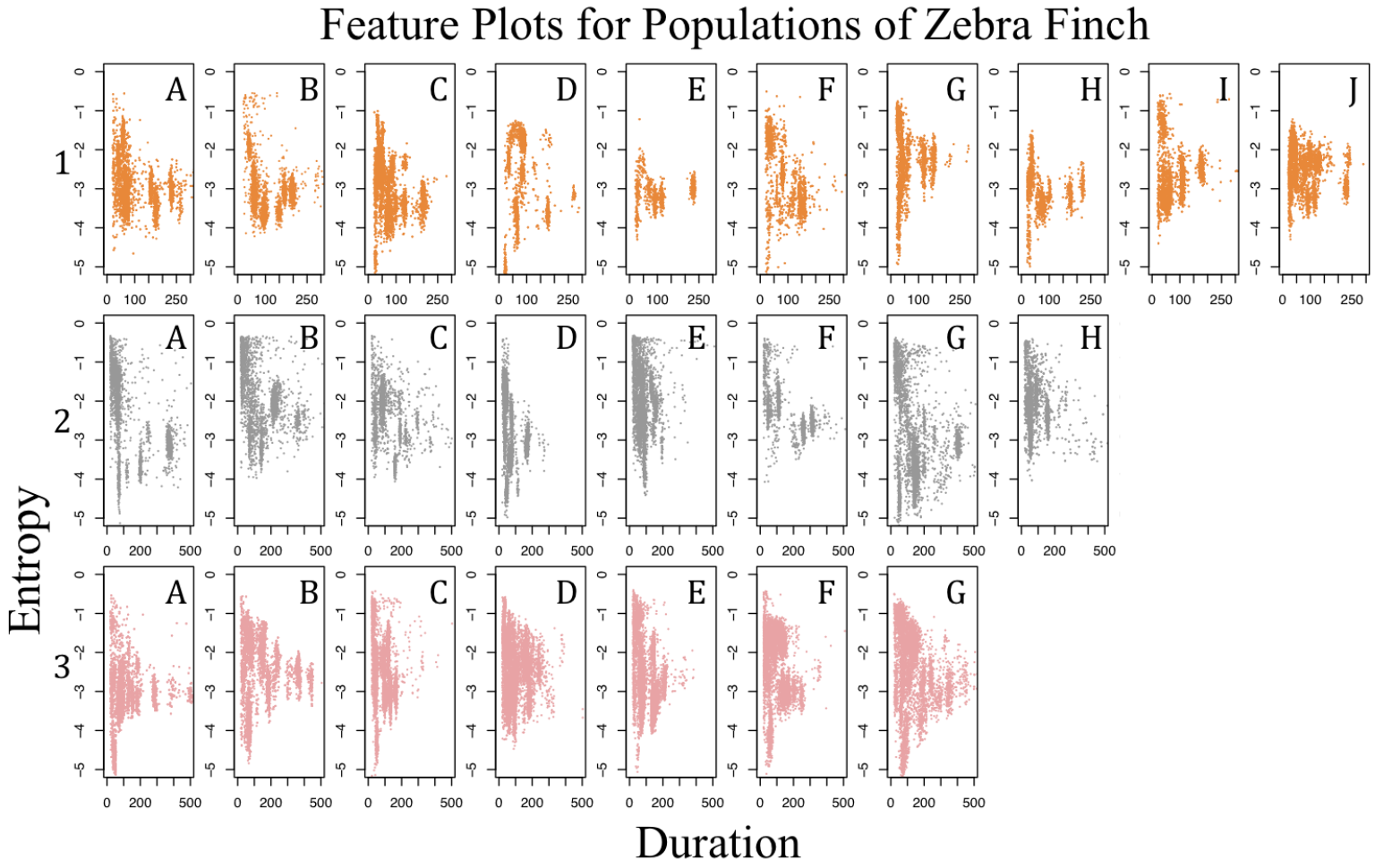

Figure 1.4. Features plots of duration vs. entropy for the 18 domesticated zebra finches and 7 wild zebra finches.

Row 1, panels A-J represent ten domesticated *T. g. castanotis* from East Carolina University. Row 2, panels A-H represent eight domesticated *T. g. castanotis* from University of Chicago. Row 3 panels A-G represent seven wild *T. g. castanotis*. Each feature plot shows duration vs. entropy. Within each parameter plot, a cluster of dots represents a note that is repeatedly sung. Therefore, the similarity in the overall pattern of clustering between each panel within a subspecies visually represents how close they are to each other.

To verify this, I ran a series of Mann-Whitney Wilcoxon tests to see if variability of song features significantly differed between populations (if the K-L distance was significantly different). For almost every comparison of every feature, there was a statistically different K-L for the two populations in question (Table 1.2). The domestic ECU *T. g. castanotis* showed significantly higher variability among individuals than wild birds for all features (Table 1.2, subset in Figure 1.5). The domestic University of Chicago *T. g. castanotis* did not show as much variability among individuals as the ECU domestic *T. g. castanotis* did. They were significantly less variable for all features. However, they still showed more variability than the wild birds for FM, entropy, variance of pitch, and variance of pitch goodness (Table 1.2, subset in Figure 1.5).

Table 1.2. Mean difference of K-L between wild and domesticated zebra finches.

Mean difference of K-L between populations (Mann-Whitney Wilcoxon test, p-values < 0.05 denoted with *). Higher K-L denotes more variability. A positive value for mean difference signifies that the top population in the comparison has higher variability. The populations for the comparison are East Carolina University domesticated *T. g. castanotis* (ECU_TGC), University of Chicago domesticated *T. g. castanotis* (UofC_TGC), and wild *T. g. castanotis* (Wild_TGC).

	Amp	Pitch	FM	AM^2	Ent.	Pitch Good.	Mean Freq	Var. Pitch	Var. FM	Var. Ent.	Var. Pitch Good.	Var. Mean Freq	Var. AM
ECU_TGC v Wild_TGC	**4.30***	**2.10***	**2.23***	**2.21***	**2.31***	**2.25***	**2.63***	**0.72***	**1.91***	**2.09***	**1.75***	**1.90***	**2.20***
UofC_TGC v Wild_TGC	**-0.26**	**0.39**	**0.79***	**0.48**	**1.19***	**0.71**	**0.09**	**0.39***	**-0.04**	**0.50**	**0.63***	**0.09**	**0.43**
ECU_TGC v UofC_TGC	**3.66***	**1.33***	**0.98***	**0.88***	**0.92***	**1.27***	**1.91***	**0.18***	**0.75***	**0.39***	**1.12***	**1.32***	**0.89***

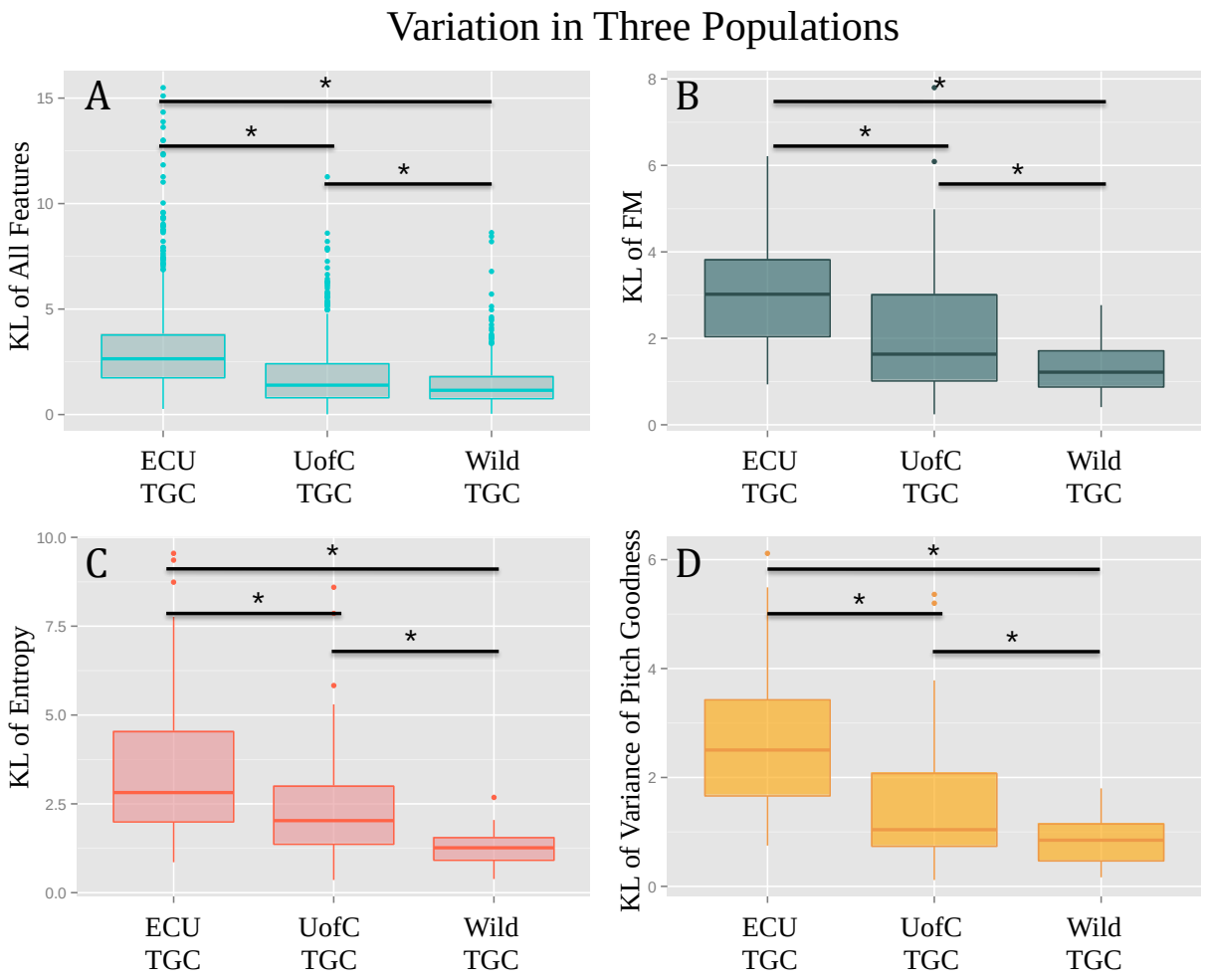

Figure 1.5. Variability (represented by K-L distance) in two populations of domestic zebra finch and one population of wild zebra finch.

K-L distance for (A) all parameters combined, (B) for FM, (C) Entropy, and (D) Pitch Goodness, of two populations of domestic *T. g. castanotis* (from East Carolina University and University of Chicago) and wild *T. g. castanotis*. Higher K-L value denotes more variation. Black bars with an * denote that there is a significant difference between the two populations it connects.

I also compared the ECU zebra finches and the population of wild zebra finches to a population of Bengalese finches. Bengalese finches are significantly less variable for all features when compared to the domesticated *T. g. castanotis* (Table 1.3). They show similar amounts of variability for 8 of the 13 features as wild *T. g. castanotis*. For the other five features, amplitude, pitch AM^2, mean frequency, and variance of AM, the Bengalese finches showed a higher amounts of variability among individuals compared to wild *T. g. castanotis* (Table 1.3).

Table 1.3. Mean difference of K-L between domestic and wild zebra finch and Bengalese finch.

Mean difference of K-L between populations (Mann-Whitney Wilcoxon test, p-values < 0.05 denoted with *). Higher K-L denotes more variability. A positive value for mean difference signifies that the top population in the comparison has higher variability. The populations for comparison are East Carolina University domesticated *T. g. castanotis* (ECU_TGC), wild *T. g. castanotis* (Wild_TGC), and East Carolina University *L. s. domestica* (ECU_LSD).

	Amp	Pitch	FM	AM^2	Ent.	Pitch Good.	Mean Freq	Var. Pitch	Var. FM	Var. Ent.	Var. Pitch Good.	Var. Mean Freq	Var. AM
ECU_LSD v ECU_TGC	-2.62*	-0.76*	-1.34*	-0.58*	-1.42*	-1.11*	-0.96*	-0.31*	-0.66*	-0.89*	-1.38*	-1.21*	-0.59*
ECU_LSD v Wild_TGC	0.79*	0.95*	0.43*	0.78*	0.69	0.87	1.04*	0.27	0.05	0.00	0.37	0.20	0.73*

Discussion

Lab models are chosen in large part due to their ability to thrive in the lab. However, drift, selection, and relaxed selection during domestication can all result in genomic changes that in turn modify behavior. This is important to consider for behavioral lab models such as the zebra finch. Especially with the importance of zebra finches as a model system, it is critical to understand the changes that are occurring as birds evolve in captivity. Here I quantified the level of divergence between captive and wild zebra finches. Using F_{ST} and diversity metrics I identified sites that vary between wild and domestic birds and found which genes these sites are near or within. I also concluded that the domesticated zebra finches show increased variability in song between individuals compared to the wild. Overall, my findings highlight genetic divergence is potentially due to both selection and drift and point to putative functional differences between wild and domesticated populations of zebra finches. My study will help us better fill in the gaps about the divergence of the domesticated birds from their wild ancestors.

Although previous studies show relatively high genetic diversity and low differentiation (Forstmeier et al. 2007), my study refines this picture. I have shown a loss of rare alleles in the domesticated zebra finch, which was also observed by Forstmeier and colleagues (2007). I find that though the populations are mildly differentiated, there are multiple signatures of selection: there is overall low levels of divergence throughout the genome, interspersed with regions of high divergence. Though an overall low genome wide F_{ST} might be mistaken as the overall subtle nature of domestication, there are multiple strong outliers that show a great deal of differentiation between wild and domesticated individuals. SNPs were near a number of potentially interesting genes, such as PRL, with functional roles that could be advantageous in domesticated environments. It is important to note that most of these highly divergent regions

were intergenic, but there is evidence that one or more could be enhancers. Deciphering these roles of these regions within this system is a future goal. Another challenge will be to improve genome assembly and annotation, utilizing the upcoming revised zebra finch genome (Korlach et al. 2017).

Although my study does not attempt to directly link genotypic and phenotypic divergence, previous studies have documented larger body size in domesticated zebra finches (Forstmeier 2007; Zann 1996) as well as longer wings, deeper bills, and taking longer to reach adult size in captive versus wild females (Zann 1996). My work adds to this by documenting differences in song behavior in selected US populations of zebra finches. Earlier studies did not find differences in syntax and phonology of song across 13 domestic populations in Europe and the United States as well as one wild population in Australia (Lachlan et al. 2016). Lachlan and colleagues (2016) did not find locally diverged cultural traditions or evidence for divergence in genetically determined constraints on song learning, despite sufficient isolation between domesticated populations to allow them to genetically diverge. They speculated that perhaps there had not been enough time for cultural evolution to take place, or, more likely in their opinion, that high error rates combined with species-constraints led to homogenization between populations (Lachlan et al. 2016). Despite these overall similarities in song structure, I have shown that domesticated zebra finches show increased variability of song features compared to wild zebra finches. One possible reason that Lachlan and colleagues (2016) did not see locally diverged cultural traditions in different populations of zebra finches was because of this increased variability the species seems to show.

The pattern of increased song variability in captive populations has also been observed in studies on the domestication of the Bengalese finch (*Lonchura striata var. domestica*) from the

white-rumped munia (*Lonchura striata*). Suzuki and colleagues (2004) found that the domesticated strains had much more complex songs, despite humans generally not artificially selecting song characteristics of this species (Suzuki et al. 2004). Suzuki and colleagues (2004) concluded that the increased complexity in Bengalese finch songs compared to the white-rumped munia initially evolved because they did not face the same environmental stress or need for species recognition. Once freed from these pressures, sexual selection then proceeded to increase the complexity of their songs (Suzuki et al. 2004). It is possible a similar scenario exists in domesticated zebra finches. However, the ECU domesticated *T. g. castanotis* showed more variability in songs between individuals than the ECU Bengalese finches, which in turn showed a similar amount of variability as wild *T. g. castanotis*. This moderate amount of variability between individuals in the wild could be because highly colonial *T. g. castanotis* occur in huge numbers in Australia, so potentially there has been selection on variability for individual recognition in the wild (within in the bounds of being able to recognize other zebra finches as conspecifics). Then, once domesticated, they possibly followed a similar trend to the Bengalese finches where when removed from the pressures faced in the wild, there was relaxed selection and increased variability in song among individuals resulted. It is possible that these domesticated birds have been freed from the constraints on song structure imposed by mate choice and the need for accurate species recognition or maybe they are behaving adaptively in a completely new and different environment compared to the wild ancestor.

Especially of interest is the finding that domestication has resulted in increased song variability among individuals. This variability has proven useful in studies of song recognition and habituation (Mello et al. 1995, Dong and Clayton 2009, Dong et al. 2009), and discrimination between a bird's own song and other conspecific songs (Chew et al. 1996; Amin

et al. 2004; Bolhuis and Gahr 2006; Bolhuis et al. 2012). Although intra-population (from the same aviary or geographic area) variation in zebra finch song seems to be on the high side of the spectrum, in wild-derived white-crowned sparrows, *Zonotrichia leucophrys*, neuronal selectivity for bird's own song has also been demonstrated (Margoliash and Konishi 1985). Likewise other species of birds can discriminate between neighbor and stranger in the wild (Brooks and Falls 1975; Lovell and Lein 2004). Thus, the high variability of domesticated zebra finch provides a rich substrate for quantifying how structural differences in song and other traits influence behavioral responses. If population differences in song variability are genetically rooted, this would facilitate genetic studies of this critical trait.

The importance of understanding genetic and phenotypic variation in model systems has already been revealed in lab mice (Yalcin et al. 2010; Casellas 2011). Inbred strains represent a single genetic identity (Yalcin et al. 2010; Nicod et al. 2016) that has been selected for under laboratory conditions, which differ vastly from natural conditions (Harper 2008). Many also contain combinations of homozygous alleles that are very rare in wild type populations (Harper 2008). Researchers have found that while inbred strains are highly useful for certain types of genetic analyses and physiological studies, where it is important to have a uniform genetic background against which to test a variant in one gene only, as well as allowed us to isolate hundreds of mutations to identify functional genes as well as isolate and maintain stem cells (Guénet and Bonhomme 2003), highly inbred strains have severe limitations for other studies such as when studying traits where it is important to understand variability among individuals. One limitation for the inbred system is when studying aging. Many inbred stocks have mutations (that cause conditions like blindness, deafness, or glaucoma) that make them incompatible with aging research (Harper 2008). Even more specifically, Dazert and colleagues (1996) found that

the pattern of age-related spiral ganglion cell degeneration is different in wild and domesticated mice. There are distinct benefits to using recently trapped wild progenitors from the taxa *Mus* for behavioral studies (Guénet and Bonhomme 2003) This is because it has been observed that many behavioral traits including fear response, activity, learning ability, reproductive behaviors, and agonistic foraging are distinctly different among subspecies and species of the genus *Mus* (Koide et al. 2000; Guénet and Bonhomme 2003). These groups can be crossed in the lab and analyzed for specific behavioral characteristics and then genotyped (Guénet and Bonhomme 2003).

Given the highly social nature of zebra finches, breeding colonies are often maintained as outbred populations, with comparatively free mating (Forstmeier et al. 2007). Diversity in captive zebra finches is still high relative to inbred lines, and probably relative to many wild species (Bulgin et al. 2003). So even though there is a loss of diversity, it is not the same level as observed in many inbred model systems (Forstmeier et al. 2007). My study is limited in that I only examined selected domesticated populations from the US, and there is evidence of distinctive ancestry in US and European birds populations (Forstmeier et al 2007). Overall it appears that drift has affected genomes and aspects of behavior domesticated zebra finch lineages. It is additionally interesting that there is likely enhanced selection on some traits, and relaxed selection on others. In general, it can appear that the effects of domestication are subtle, but I have found highly diverged aspects of the genome and song production. My study suggests there is potentially a lot to be gained by taking characterizing and taking advantage of the full spectrum of genetic and behavioral diversity in this species.

References

Alachiotis N, Stamatakis A, Pavlidis P (2012) OmegaPlus: a scalable tool for rapid detection of selective sweeps in whole-genome datasets. *Bioinformatics*, 28:2274–2275.

Amin N, Grace JA, Theunissen FE (2004) Neural response to bird's own song and tutor song in the zebra finch field L and caudal mesopallium. *Journal of Comparative Physiology A, 190*(6): 469-489.

Angelier F, Chastel O (2009) Stress, prolactin and parental investment in birds: A review. *Gen. Comp. Endocrinol*, 163:142–148.

Balakrishnan CN, Edwards SV (2009) Nucleotide variation, linkage disequilibrium and founder-facilitated speciation in wild populations of the zebra finch (*Taeniopygia guttata*). *Genetics*, 181: 645-660.

Benítez-Burraco A, Theofanopoulou C, Boeckx C (2016) Globularization and Domestication. *Topoi*, 1-14.

Bolhuis JJ, Gahr M (2006) Neural mechanisms of birdsong memory. *Nature reviews neuroscience*, *7*(5):347-357.

Bolhuis JJ, Gobes SM, Terpstra NJ, den Boer-Visser AM, Zandbergen MA (2012) Learning-related neuronal activation in the zebra finch song system nucleus HVC in response to the bird's own song. *PloS one*, *7*(7):e41556.

Brooks RJ, Falls JB (1975) Individual recognition by song in white-throated sparrows. I. Discrimination of songs of neighbors and strangers. *Canadian Journal of Zoology*, *53*(7):879-888.

Bulgin NL, Gibbs HL, Vickery P, Baker AJ (2003) Ancestral polymorphisms in genetic markers obscure detection of evolutionarily distinct populations in the endangered Florida grasshopper sparrow (Ammodramus savannarum floridanus). *Mol Ecol*, 12(4):831–844.

Chen H, Patterson N, Reich D (2010) Population differentiation as a test for selective sweeps. *Genome Res*., 20(3):393–402.

Chew SJ, Vicario DS, Nottebohm F (1996) A large-capacity memory system that recognizes the calls and songs of individual birds. *Proceedings of the National Academy of Sciences*, *93*(5): 1950-1955.

Clayton NS (1990a) Assortative Mating in Zebra Finch Subspecies, Taeniopygia guttata guttata and *T. g. castanotis*. *Philosophical Transactions: Biological Sciences*, 330(1258),:351-370.

Clayton NS (1990b) Mate choice and pair formation in Timor and Australian Mainland zebra finches. *Anim. Behav*., 39:474~480.

Clayton NS (1990c) Subspecies recognition and song learning in zebra finches. *Anim. Behav*., 40:1009-1017.

Clayton NS (1989) The effects of cross-fostering on selective song learning in estrildid finches. *Behaviour,* 190:163-74.

Dave AS, Margoliash D (2000) Song Replay During Sleep and Computational Rules for Sensorimotor Vocal Learning. *Science,* 290(5492):812-816.

Dazert S, Feldman ML, Keithley EM (1996) Cochlear spiral ganglion cell degeneration in wild-caught mice as a function of age. *Hearing Res,* 100:101–106.

Deacon TW (2010) Colloquium paper: a role for relaxed selection in the evolution of the language capacity. *Proc Natl Acad Sci USA,* 107:9000–9006.

Derégnaucourt, S., P. P. Mitra, O. Fehér, C. Pytte, and O. Tchernichovski. 2005. How sleep affects the developmental learning of bird song. *Nature,* 433:710–6.

Dong S, Clayton D F (2009) Habituation in songbirds. *Neurobiology of learning and memory, 92*(2):183-188.

Dong S, Replogle KL, Hasadsri L, Imai BS, Yau PM, Rodriguez-Zas S, ... & Clayton DF (2009) Discrete molecular states in the brain accompany changing responses to a vocal signal. *Proceedings of the National Academy of Sciences*, 106(27):11364-11369.

ElemeNT -- http://lifefaculty.biu.ac.il/gershon-tamar/index.php/element-description/element

Estep DQ, Lanier DL, Dewsbury DA (1975) Copulatory behavior and nest building behavior of wild house mice (Mus musculus). *Anim Learn Behav,* 3:329–336.

Forstmeier W, Segelbacher G, Mueller JC, Kempenaers B (2007) Genetic variation and differentiation in captive and wild zebra finches (*Taeniopygia guttata*). *Molecular ecology* 16:4039–50.

Forstmeier W, Burger C, Temnow K, Derégnaucourt S (2009) The genetic basis of zebra finch vocalizations. *Evolution,* 63:2114-2130.

Frank H, Frank MG (1982) On the effects of domestication on canine social development and behavior. *Appl Anim Ethol,* 8:507–525.

Gilby AJ, Mainwaring MC, Griffith SC (2013) Incubation behaviour and hatching synchrony differ in wild and captive populations of the zebra finch. *Anim Behav, 85*(6):1329-1334.

Gilby AJ, Mainwaring MC, Rollins LA, Griffith SC (2011) Parental care in wild and captive zebra finches: Measuring food delivery to quantify parental effort. *Anim Behav,* 81:289–295.

Guénet JL, Bonhomme F (2003) Wild mice: an ever-increasing contribution to a popular mammalian model. *Trends in Genetics, 19*(1):24-31.

Harper JM (2008) Wild-derived mouse stocks: An underappreciated tool for aging research. *Age (Omaha),* 30:135–145.

Hare B, Brown M, Williamson C, Tomasello M (2002) The domestication of social cognition in dogs. *Science, 298*(5598):1634-1636.

Holveck MJ, Vieira de Castro AC, Lachlan RF, ten Cate C, Riebel K (2008) Accuracy of song syntax learning and singing consistency signal early condition in zebra finches. *Behav Ecol, 19*(6):1267-1281.

Houx AB, ten Cate C (1999) Song learning from playback in zebra finches: is there an effect of operant contingency? *Anim Behav, 57*(4)837-845.

Jin H, Clayton DF (1997) Localized Changes in Immediate-Early Gene Regulation during Sensory and Motor Learning in Zebra Finches. *Neuron,* 19:1049–1059.

Jones AE, Ten Cate C, Slater PJ (1996) Early experience and plasticity of song in adult male zebra finches (Taeniopygia guttata). *Journal of Comparative Psychology,* 110(4):354.

Kardong K (1993) The predatory behavior of the Northern Pacific rattlesnake (*Crotalus viridis oreganus*): laboratory versus wild mice as prey. *Herpetologica,* 49:457–463.

Koide T, Moriwaki K, Ikeda K, Niki H, Shiroishi T (2000) Multi-phenotype behavioral characterization of inbred strains derived from wild stocks of Mus musculus. *Mammalian Genome,* 11(8):664-670.

Korneliussen TS, Albrechtsen A, Nielsen R (2014) ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinformatics,* 15(1):356.

Korneliussen TS, Moltke I, Albrechtsen A, Nielsen R (2013) Calculation of Tajima's D and other neutrality test statistics from low depth next-generation sequencing data. *BMC Bioinformatics,* 14:289.

Korlach J, Gedman G, Kingan S, Chin J, Howard J, Cantin L, Jarvis ED (2017) De Novo PacBio Long-Read and Phased Avian Genome Assemblies Correct and Add to Genes Important in Neuroscience Research. *bioRxiv*, http://biorxiv.org/content/early/2017/02/02/103911.abstract.

Kronenberger J-P, Medioni J (1985) Food neophobia in wild and laboratory mice (*Mus musculus domesticus*). *Behav. Processes,* 11(1):53–59.

Lachlan RF, van Heijningen CA, Ter Haar SM, Ten Cate C (2016) Zebra Finch song phonology and syntactical structure across populations and continents—a computational comparison. *Frontiers in Psychology, 7*.

Lacy RC (1997) Importance of genetic variation to the viability of mammalian populations. *Journal of Mammalogy,* 78(2):320-335.

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler Transform. *Bioinformatics,* 25:1754-60.

Librado P, Der Sarkissian C, Ermini L, Schubert M, Jónsson H, Albrechtsen A, ... & Mortensen CD (2015) Tracking the origins of Yakutian horses and the genetic basis for their fast adaptation to arctic environments. *Proc Natl Acad Sci USA, 112*(50):E6889-E6897.

London SE, Clayton DF (2008) Functional identification of sensory mechanisms required for developmental song learning. *Nature Neuroscience,* 11:579-586.

Lovell SF, Lein MR (2004). Neighbor-stranger discrimination by song in a suboscine bird, the alder flycatcher, Empidonax alnorum. *Behavioral Ecology, 15*(5):799-804.

Mainwaring MC, Hartley IR, Gilby AJ, Griffith SC (2010) Hatching asynchrony and growth trade offs within domesticated and wild zebra finch broods. *Biol J Linn Soc,* 100:763–773.

Mann NI, Slater PJB (1995) Song tutor choice by zebra finches in aviaries. *Animal Behaviour,* 49:811-820.

Mariette MM, Griffith SC (2012) Nest visit synchrony is high and correlates with reproductive success in the wild Zebra finch Taeniopygia guttata. *J Avian Biol,* 43:131–140.

Marler P (1970) Birdsong and speech development: could there be parallels? *American Scientist,* 58:669-73.

Mayr E (1944) Timor and the colonization of Australia by birds. *Emu,* 44:113-30.

Mello C, Nottebohm F, Clayton D (1995) Repeated exposure to one song leads to a rapid and persistent decline in an immediate early gene's response to that song in zebra finch telencephalon. *J Neurosci,* 15:6919–25.

Nicod J, Davies RW, Cai N, et al. (2016) Genome-wide association of multiple complex traits in outbred mice by ultra-low-coverage sequencing. *Nat Genet,* 48(8):912–918.

Mooney R (2009) Neural mechanisms for learned birdsong. *Learn Mem,*16:655–69.

Newhouse DJ, Balakrishnan CN (2015) High MHC diversity despite bottlenecks in wild and domesticated zebra finches. *BMC Evolutionary Biology,*15:256.

NNPP -- http://www.fruitfly.org/seq_tools/promoter.html

Olveczky BP, Andalman AS, Fee MS (2005) Vocal experimentation in the juvenile songbird requires a basal ganglia circuit. *PLoS Biology,* 3:e153.

Poh Y-P, Domingues VS, Hoekstra HE, Jensen JD (2014) On the prospect of identifying adaptive loci in recently bottlenecked populations. *PloS one,* 9:e110579.

Price EO (1999) Behavioral development in animals undergoing domestication. *Applied Animal Behaviour Science*, *65*(3):245-271.

Price T (2008) *Speciation in birds*. Roberts and Company Publishers.

R Core Team (2014) R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

RAVEN (Bioacoustics Research Program 2011).

Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP (2011) Integrative Genomics Viewer. *Nature Biotechnology,* 29:24–26.

Rogers CH (1979) Zebra finches. K & R Books.

Rubin CJ et al. (2010) Whole-genome resequencing reveals loci under selection during chicken domestication. *Nature,* 464:587–91.

Rutstein AN, Brazill-Boast J, Griffith SC (2007) Evaluating mate choice in the zebra finch. *Anim Behav,* 74:1277–1284.

Schütz KE, Forkman B, Jensen P (2001) Domestication effects on foraging strategy, social behaviour and different fear responses: A comparison between the red junglefowl (*Gallus gallus*) and a modern layer strain. *Appl Anim Behav Sci,* 74:1–14.

Schütz KE, Jensen P (2001) Effects of resource allocation on behavioural strategies: a comparison of red junglefowl (Gallus gallus) and two domesticated breeds of poultry. *Ethology*, *107*(8):753-765.

Shapiro MD, Kronenberg K, Li C, Domyan ET, Pan H, Campbell M, Tan H, Huff CD, Haofu Hu, Vickrey AI, Nielsen SA, Stringham SA, Hao Hu, Willerslev E, Gilbert MTP, Yandell M, Zhang G, Wang J (2013) Genomic diversity and evolution of the head crest in the rock pigeon. *Science,* 339:1063-1067.

Singhal S, Leffler EM, Sannareddy K, Turner I, Venn O, Hooper DM, Strand AI, Li Q, Raney B, Balakrishnan CN, Griffith SC, McVean G, Przeworski M (2015) Stable recombination hotspots in birds. *Science,* 350(6263):928-932.

Slater PJB, Eales LA, Clayton NS (1988) Song learning in zebra finches (Taeniopygia guttata): progress and prospects. *Advances in the Study of Behavior,* 18:1-34.

Smith J, Hurst JL, Barnard CJ (1994) Comparing behaviour in wild and laboratory strains of the house mouse: levels of comparison and functional inference. *Behav. Processes*, 32:79–86.

Sturdy CB, Phillmore LS, Weisman RG (1999) Note types, harmonic structure, and note order in the songs of zebra finches (*Taeniopygia guttata*). *Journal of Comparative Psychology*, 113(2):194.

Suzuki K, Ikebuchi M, Bischof H, Okanoyaa K (2014) Behavioral and neural trade-offs between song complexity and stress reaction in a wild and a domesticated finch strain. *Neuroscience and Biobehavioral Reviews*, 46:547–556.

Tchernichovski O, Nottebohm F, Ho CE, Pesaran B, Mitra PP (2000) A procedure for an automated measurement of song similarity. *Anim Behav*, 59:1167–1176.

ten Cate C (2014) On the phonetic and syntactic processing abilities of birds: From songs to speech and artificial grammars. *Current Opinion in Neurobiology*, 28:157–164.

Thompson JA, Basista MJ, Wu W, Bertram R, Johnson F (2011) Dual Pre-Motor Contribution to Songbird Syllable Variation. *The Journal of Neuroscience*, 31(1):322–330.

Thorvaldsdóttir H, Robinson JT, Mesirov JP (2013) Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in Bioinformatics*, 14:178-192.

Trut L, Oskina I, Kharlamova A (2009) Animal evolution during domestication: The domesticated fox as a model. *BioEssays*, 31:349–360.

Tsutsui ND, Suarez AV, Holway DA, Case TJ (2000) Reduced genetic variation and the success of an invasive species. *Proceedings of the National Academy of Sciences*, *97*(11):5948-5953.

Vallentin D, Kosche G, Lipkind D, Long MA (2016) Inhibition protects acquired song segments during vocal learning in zebra finches. *Science* 351(6270):267-271.

vcflib – https://github.com/vcflib/vcflib

Vickrey AI, Domyan ET, Horvath MP, Shapiro MD (2015) Convergent Evolution of Head Crests in Two Domesticated Columbids Is Associated with Different Missense Mutations in EphB2. *Molecular biology and evolution*, 32:2657-2664.

Warren WC, Clayton DF, Ellegren H, Arnold AP, Hillier LW, Künstner A, Searle S, White S, Vilella AJ, Fairley S et al. (2010) The genome of a songbird. *Nature*, 464(7289):757-62.

Woolley SC, Doupe AJ (2008) Social context-induced song variation affects female behavior and gene expression. *PLoSBiol*, 6(3):e62.

Wu W, Thompson JA, Bertram R, Johnson F (2008) A statistical method for quantifying songbird phonology and syntax. *Journal of Neuroscience Methods,* 174:147–154.

Zann RA (1996) *The zebra finch: a synthesis of field and laboratory studies*. Oxford: Oxford University Press.

Zann RA (1993) Variation in song structure within and among populations of Australian zebra finches. *The Auk,* 110(4):716-726.

CHAPTER 2 GENETIC AND BEHAVIORAL DIVERGENCE DURING SPECIATION IN THE ZEBRA FINCH (TAENIOPYGIA GUTTATA)

Abstract

With their learned vocalization serving as an ideal model of human speech acquisition, the zebra finch, *Taeniopygia guttata*, is a classic neurological and behavioral system. With the sequencing of its genome it is also rising in importance in the field of population genetics. Zebra finches are diverging in allopatry without secondary contact and the intraspecific variation between the two subspecies has not been well characterized. In fact much of the focus remains on one of the two subspecies, ignoring the potential to investigate the genotypic and behavioral variation that has resulted due to allopatric speciation. In this study I seek to characterize variation in both of these regards. I investigated if the previously documented reduction of diversity in a subset of noncoding loci is present at the genome level. I found further support for a dramatic bottleneck in the island birds as the two subspecies have diverged, as evidenced by an overall reduction in diversity. I also focused on two phenotypic traits, pigmentation and vocal behavior, and found differences between the two subspecies. I have identified candidate SNPs in two well-known pigmentation genes, SLC45A2 and CDKN2A, and have found that the island subspecies shows less variation in song among individuals than the mainland birds. Overall, I have characterized the genetic and behavioral variation in the zebra finch system, which has the potential to further our ability to study behaviors such as heritability of song learning.

Introduction

As a representative of the Oscine Passerines ("songbirds"), birds with learned vocalizations, the zebra finch, *Taeniopygia guttata*, has been the subject of wide-ranging behavioral and neurological research (Jin and Clayton 1997; Dave and Margoliash 2000; Olveczky et al. 2005; Mooney 2009; London and Clayton 2008; Thompson et al. 2011; Vallentin et al. 2016). Much of the research done on zebra finches is centered on vocal learning behavior (Forstmeier et al. 2009; reviewed in Slater et al. 1988), a trait that songbirds share with humans (Mello 2014; Marler 1970; ten Cate 2014). Now that there is a complete zebra finch genome (Warren et al. 2010), the zebra finch is now also a rising model system for genomics (Clayton 2004). Despite the fact that zebra finches are maintained as outbred study organisms (Forstmeier et al. 2007), research to date has largely ignored the genetic differences between the two subspecies, *Taeniopygia guttata castanotis* (*T. g. castanotis*) and *Taeniopygia guttata guttata* (*T. g. guttata*), and the possibility of intraspecific variation in the learning and production of song.

Another underutilized resource of this system is the chance to investigate an ongoing divergence in allopatry without secondary contact. As speciation research has shifted to emphasize "complex speciation," in many ways we lack a simple null model of genomic divergence in the absence of gene flow, and the zebra finch subspecies represent just that scenario. Allopatric speciation with secondary contact has been well-described (Hoskin et al. 2005; Kuehne et al. 2007; Grant and Grant 2009). Divergence without secondary contact is less studied. In this study I am interested if divergence is uniform in allopatry. In a situation where gene exchange is impossible, reproductive isolation can evolve in a number of ways: sexual selection, uniform selection, stabilizing selection, or divergent selection. And it can happen quickly such as with divergent selection or slowly under balancing or uniform selection (Via

2009). I seek to test the generality of these novel findings in a vertebrate model system with a well-defined biogeographic history (Balakrishnan and Edwards 2009).

Indeed most research to date is focused on only one of the two subspecies, the Australian zebra finch (*T. g. castanotis*) and has used domesticated strains of this subspecies, which have been bred in captivity and well-studied in the laboratory for over 150 years. They are extremely common in the wild and are distributed all across the continent except in the far South and North (Zann 1996). A second subspecies, the Timor zebra finch (*T. g. guttata*), native to the Lesser Sunda Islands north of the continent, has received minimal attention (e.g., Clayton 1990a-c). Behavioral differences between the two subspecies have been characterized (Clayton 1990; Clayton et al. 1991), but their divergence is less well understood (but see Balakrishnan and Edwards 2009). Currently, the Lesser Sunda Islands are 300 miles from the closest point on the Australian continent. However, this distance was far smaller (estimated 45-75 miles) in recent geological history due to glaciation events when lower sea levels may have facilitated the crossing between Australia and the Lesser Sundas. It is estimated that 20-22 species of bird traveled from the Lesser Sundas to Australia during this period. During this period some species went in the opposite direction, including the zebra finch (Mayr 1944). Since their colonization 1-2 MYA (Balakrishnan and Edwards 2009), the two populations have diverged in allopatry. In this particular system the two subspecies will interbreed in captivity and produce viable, fertile offspring (personal observation). This biogeographic context of an ongoing speciation event allows for an interesting opportunity to investigate how those natural processes have impacted both genetic and behavioral variation.

A previous study by Balakrishnan and Edwards (2009) found that for 30 noncoding loci the *T. g. guttata* subspecies showed a comparative lack of genetic diversity. This suggests that

the colonization of the islands involved a population bottleneck (Balakrishnan and Edwards 2009) where there was a sudden reduction in population size, which can occur when a new population is founded (Coyne and Orr 2004). This can result in a loss of genetic variation, especially the loss of rare alleles in the small founder population, since they would only carry a proportion of the variability from the parent population (Mayr 1942). Selection pressures will differ on the founder population since alleles will be selected to function on a more homozygous background as opposed to the parental heterozygous background (Coyne and Orr 2004). One goal of this study is to investigate if this reduction in diversity within the *T. g. guttata* species found by Balakrishnan and Edwards (2009) is present at the genome-wide level as well as characterize aspects of genotypic and behavioral variation.

Along with further examining the genomic impact of this speciation event, I also wanted to look into two traits that might have resulted due to allopatric speciation of the zebra finch and see if some well-documented trends of island vs. mainland bird species were found in this system: pigmentation and vocal behavior. Plumage of island birds is typically duller and dimorphism between the sexes is reduced. One explanation for this difference is due to varying ecological conditions (Price 2008). It has already been documented that zebra finches fit these patterns with reduced plumage color and body size dimorphism. The island *T. g. guttata* are statistically smaller that *T. g. castanotis* (Forstmeier et al. 2007) and they lack the black barring along the upper breast, which is characteristic of mainland birds (Clayton et al. 1991; Zann 1996). I preliminarily investigated two known pigmentation genes, SLC45A2 and CDKN2A, as possible candidate SNPs for impacting the coloration differences between the two subspecies. CDKN2A in particular is another known pigmentation gene that has been linked to barring in chickens (Hellström et al. 2010). Both SLC45A2 and CDKN2A are located on the Z

chromosome in birds. The vocal behavior trend that I focused on is the trend that some species, such as reed warblers (*Acrocephalus* spp.), show where the island species have songs that are less complex than the mainland species (Catchpole and Komdeur 1993). One justification for this is that these changes are due to cultural drift. Often it is a small population that colonizes an island, which leads to founder effects resulting in reduced syllable diversity (Price 2008). Another hypothesis is that sexual selection is predicted to be a weaker force on islands when compared to mainland populations, which has been demonstrated by less extra pair copulation in socially monogamous island passerines (Griffith 2000), and this could result in less complex songs (Price 2008). Additionally, Verner (1964) argues that lower breeding densities due to harsh island conditions means that all surviving males are in relatively high quality areas. Therefore polygyny does not evolve as readily because there is lower variation in resource quality among territories (Verner 1964). The proposed explanation for less complex songs is that if songs are not as important for securing mates and extra pair copulations, then there would be less of a drive for the males' songs to be complex if females are not strongly selecting for complexity.

Overall, despite the importance of zebra finches as a model system for social behavior and vocal learning, there is an important gap in our knowledge of how genomes and behavior vary among populations. This study is the start to characterizing that variation. I find further evidence for a bottleneck causing the reduction of diversity in *T. g. guttata* across the genome. I find that there is a reduction in the variability of songs among individuals in *T. g. guttata*. And lastly, I identify potential SNPs in pigmentation genes that could have a role in the plumage differences between the subspecies.

Methods

Genomic divergence

Sample collection and sequencing. In order to derive a complete picture of genetic variation, a total of 28 unrelated individuals (19 wild *T. g. castanotis* and 9 *T. g. guttata*) were sequenced. Three of the nine *T. g. guttata* were sampled from our laboratory populations at East Carolina University. All three were males. Permissions were obtained for the samples and they were collected under the supervision of the relevant ethical committees. The remaining six *T. g. guttata* came from museum samples from Commonwealth Scientific and Industrial Research Organisation (CSIRO) in Australia. These specimens are around 50 years old collected from Baucau Province in East Timor. Researchers such as McCormack and colleagues (2016) have had success applying next generation sequencing methods to target ultraconserved elements in museum samples of western scrub jays ranging back 120 years despite older samples yielding fewer and shorter loci in general. Hung and colleagues (2014) did full genome sequencing on museum specimen of passenger pigeon toe pads, obtaining sequencing depths of 5-20x, but there have not been many studies that have used museum toe pads for next-gen sequencing since these preserved tissues yield little DNA (Burrell et al. 2015). Of the six museum samples, four were males and two were females. For each wild zebra finch, genomic DNA was extracted using a DNeasy Blood and Tissue Kit (Qiagen kit cat no. 69504). Paired-end sequencing libraries were constructed following Illumina protocol for sequencing on the Illumina HiSeq 2000 platform. This platform was used to re-sequence paired end full genomes at medium coverage (8x).

Nineteen *T. g. castanotis* genomes were sequenced using the Illumina HiSeq platform at medium coverage (20x) by collaborators Singhal et al. (2015). The *T. g. castanotis* birds (10 females and 9 males) were collected from a studied population from Fowlers Gap in New South

Wales, Australia (Singhal et al. 2015). From the one *T. g. castanotis* birds, DNA was extracted from blood using a DNeasy Blood and Tissue Kit (Qiagen kit cat no. 69504). A species pooled library was diluted to 10nM and sequenced across 15 lanes using 100 paired end reads with an Illumina HiSeq 2000. This was done in 2012 at the Oxford Genomics Centre at the Wellcome Trust Centre for Human Genetics in Oxford, UK. For the two domesticated birds, DNA was extracted using a DNeasy Blood and Tissue Kit (Qiagen kit cat no. 69504). A pooled library was made and sequenced across six lanes on 2-lane flow cells of an Illumina HiSeq using 100 base pair paired-end reads. This was done in 2013 at the University of Chicago Genomics Core, Chicago IL USA (Singhal et al. 2015).

Steps for quality checking, mapping, and downstream analysis using ANGSD to call SNPs and look at metrics for assessing diversity followed the same methods utilized in chapter 1.

Screening for outlier F_{ST} sites. As with the domestication analysis in chapter 1, I used the SegmentF_{ST} method in vcflib to find continuous regions with high F_{ST} values (vcflib – https://github.com/vcflib/vcflib). The tool SegmentF_{ST} scans the output of wcF_{ST} that I obtained from ANGSD and measures the number of F_{ST} values that are above the threshold I set at 0.9. As the chapter 1 methods describe, here I also used the tool permuteSmooth and obtained outlier sites of high F_{ST}.

Pigmentation genes

Sequencing of SLC45A2. Genomic DNA was extracted from fourteen individuals for the targeted approach of investigating SLC45A2. Three *T. g. guttata* males from East Carolina University that were used for investigating genomic divergence were also used for this portion.

Eleven *T. g. castanotis* were also used, three of which were wild type coloration and eight of which had various color morphs that are commonly selected for by zebra finch breeders and enthusiasts. Seven exons of SLC45A2 were amplified using polymerase chain reaction. After PCR products were cleaned they were sequenced at East Carolina University and these data were then loaded into *Geneious* (http://www.geneious.com, Kearse et al., 2012) to look for possible SNPs between the morphs. Any SNPs recovered that differed between *T. g. castanotis* and *T. g. guttata* individuals were investigated further.

Investigation of the pigmentation genes on the Z chromosome. After read-mapping, the sorted bam alignment files from the nineteen *T. g. castanotis* and the nine *T. g. guttata* individuals were loaded into the Integrative Genome Viewer (IGV) (Robinson et al. 2011; Thorvaldsdóttir et al. 2013) using the long tail finch genome as the reference. The SNP identified in *Geneious* after sequencing SLC45A2 was investigated to see if it was derived in the *T. g. guttata* birds relative to *T. g. castanotis* and *P. acuticauda.*

I also used IGV to scan the CDKN2A gene to see if I could identify other SNPs in this other well-studied pigmentation gene. Because I had not done any targeted PCR in this region, I manually scanned the whole sequence between chrZ:32,024,669-32,039,744 looking for looking for single base pair polymorphisms.

Analysis of song variability

In order to characterize the variation in the patterns of song behavior between different populations I analyzed songs from a wild *T. g. castanotis* population to compare to a *T. g. guttata*

population. Ten *T. g. guttata* birds were recorded at East Carolina University and seven wild *T. g. castanotis* were recorded at Macquarie University in Sydney, Australia.

Study Populations

The *T. g. guttata* colony was originally founded at the University of Illinois with five pairs of birds. The descendants of this founding colony (~30 birds) were moved to East Carolina University in 2012. Roy Beckham brought the *T. g. guttata* into captivity 25 years ago (efinch.com). All birds were housed in large flight aviaries at the Brody School of Medicine where they could freely feed on a commercial finch dry-seed mix, with fresh water provided daily and warmed, finely crushed hard-boiled eggs and greens given weekly. The birds were monitored on a daily basis. The East Carolina University Animal Ethics Committee approved housing and experimental protocols.

The seven wild *T. g. castanotis* are birds that were taken from the wild in 2007 from Sturt National Park in far northwest New South Wales and were allowed to breed in captivity in a couple of large aviaries. In 2010 a further set of birds were taken from Fowlers Gap (far-west NSW) and added to the population in captivity. In total about 100 adults were taken from Sturt with about 40 adults taken from Fowlers Gap. These populations housed by my collaborator Simon Griffith in his lab at Macquarie University in Sydney, Australia. They have been held in isolation (from domesticated birds) since then and have mostly had the opportunity to breed about every 12 months (Gilby et al. 2013). Chris Balakrishnan recorded them using the same method utilized for the birds at ECU in the summer of 2016.

Song Recordings

The song recording methods followed the same steps as laid out in chapter 1.

Statistical Analyses

The statistical analyses methods followed the same steps as laid out in chapter 1.

Results

Genomic Divergence

Genomic polymorphism & divergence in the two subspecies of finches. As might be expected, diversity is reduced in the island subspecies: diversity in *T. g. guttata* is reduced 97.6% from *T. g. castanotis* ($p<0.001$). Average *theta* for *T. g. castanotis* was 15.53 whereas average *theta* for *T. g. guttata* was 3.15. This difference is consistent across the genome with 99% of 1000bp windows showing lower diversity in *T. g. guttata* versus *T. g. castanotis*. The Watterson Estimator, *theta*, estimated by the number of polymorphic sites, is the product of effective population size and the neutral mutation rate to show overall population mutation rate. Therefore, a higher value for *theta* is associated with greater diversity. Average Tajima's D also differed significantly between populations (*T. g. castanotis* = -1.515, *T. g. guttata* = -1.103, 0.22% less negative in *T. g. guttata* at 69% of the sites $p<0.01$). The value for Tajima's D becomes more negative when there is an excess of rare alleles. Along with the overall loss in diversity, there has been a loss of rare alleles, seen in the shift in the SFS (Figure 2.1). *T. g. castanotis* have a higher proportion of alleles that are only represented once and show an overall variability at 7.47% of the sites whereas *T. g. guttata* show an overall variability of 1.05% of the sites (Figure 2.1).

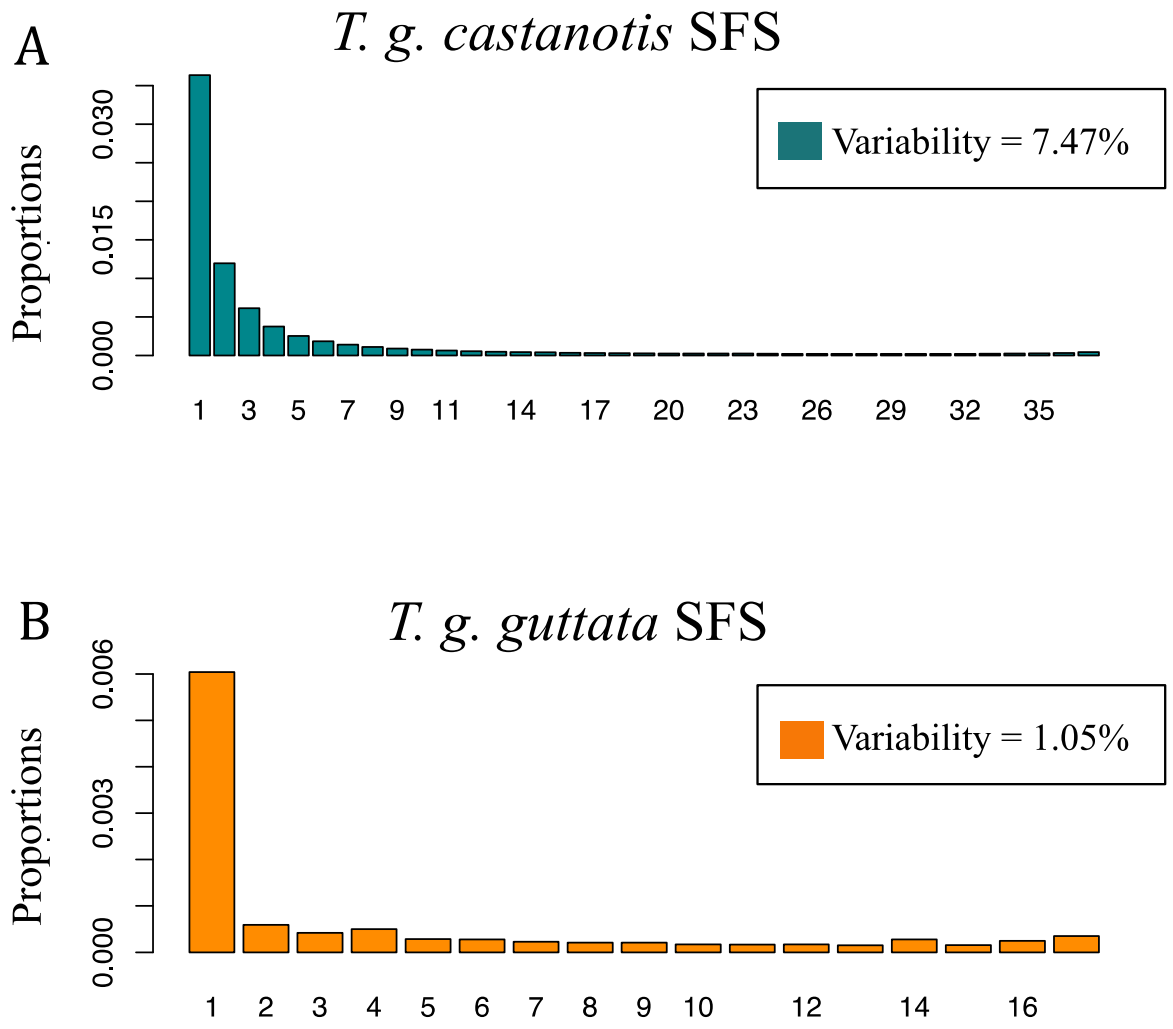

Figure 2.1. Site frequency spectra for wild and domesticated zebra finches.

T. g. castanotis show a relative excess of rare alleles. The number of bins on the X-axis is equal to 1 minus the number of alleles in each population (37 for the 19 *T. g. castanotis* and 16 for the 9 *T. g. guttata*). The proportion represents the proportion alleles are present one, or two, or three and more times. Variability represents the percent of the sites that are variable.

I find that populations show a low level of overall genetic differentiation (F_{ST} = 0.001), but in the sliding windows approach, F_{ST} ranged up to 1, indicating regions of high differentiation.

Selection or drift during speciation. Despite overall low F_{ST}, it is possible that drift or selection has led to divergence in specific areas of the genome. I used multiple approaches to examine selection. I identified 45801 outlier regions of high F_{ST} using the program SegmentFst from vcflib (Figure 2.2).

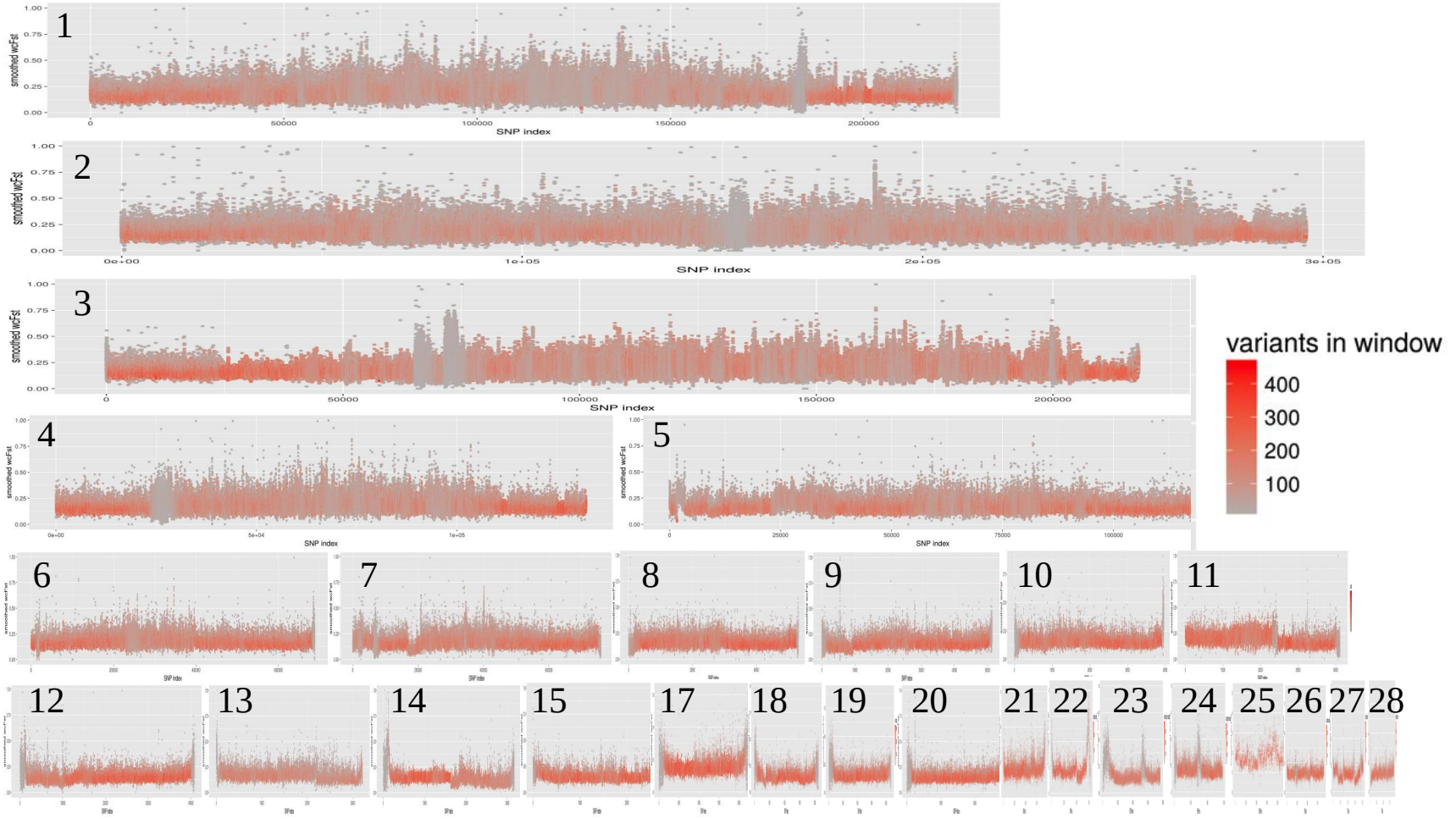

Figure 2.2. Windowed F_{ST} by chromosome.

Smoothed regions of F_{ST} using windows of 2000 base pairs and step of 500 base pairs broken down by chromosome

Pigmentation candidate genes

SLC45A2

From the PCR method, a SNP was identified on exon 7 of SLC45A2 in the three *T. g. guttata* males. These individuals carry a thymine nucleotide compared to the cytosine found in the other eleven birds (Figure 2.3).

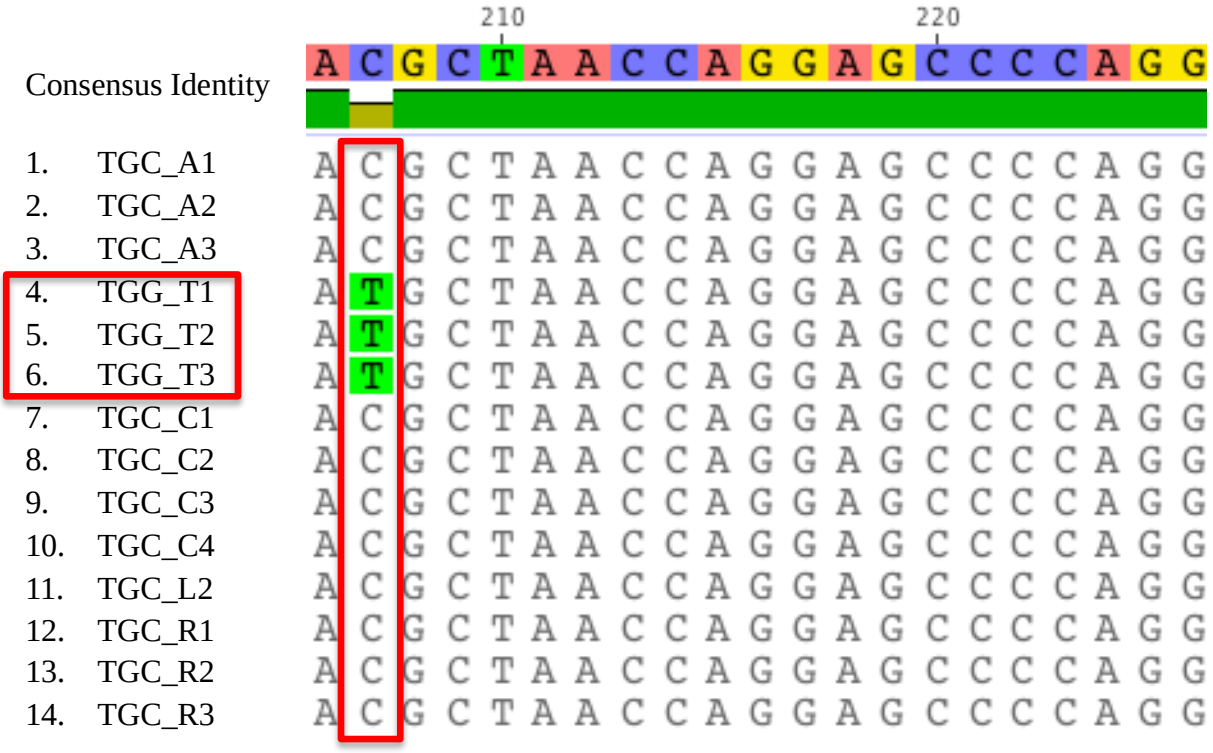

Figure 2.3. Nucleotide sequence for a portion Exon seven of SLC45A2 for three *T. g. guttata* and eleven *T. g. castanotis.*

Nucleotide substitution from C to T seen in the three male *T. g. guttata* in exon 7 of SLC45A2.

This C-T substitution yielded an amino acid change from valine in the *T. g. castanotis* individuals to isoleucine in the *T. g. guttata* individuals (Figure 2.4).

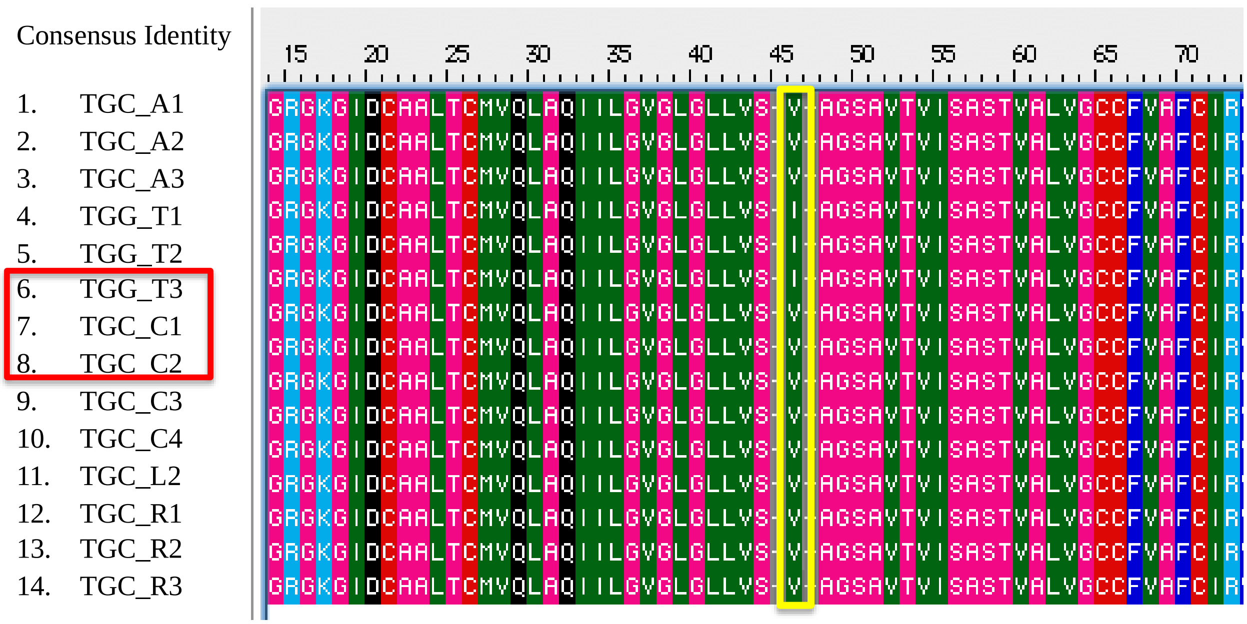

Figure 2.4. Amino acid sequence for a portion of exon seven of SLC45A2 for three *T. g. guttata* and eleven *T. g. castanotis.*

Amino acid substitution from V to I seen in the three male *T. g. guttata* in exon 7 of SLC45A2.

When I looked at this same region using the bam files of the 19 wild *T. g. castanotis* and the 9 *T. g. guttata* this polymorphism was confirmed and it appears to be derived in *T. g. guttata* birds. The F_{ST} for this site is 0.99.

CDKN2A

After manually scanning the whole sequence between chrZ:32,024,669-32,039,744 looking for any instances where one base pair was derived in the *T. g. guttata* birds. I found six SNPs where this was the case (Table 2.1).

Table 2.1. Position and nucleotide substitution for six sites in CDKN2A for *T. g. castanotis* and *T. g. guttata*.

Six instances of nucleotide substitutions seen in *T. g. guttata* vs. *T. g. castanotis* birds and their position along the Z chromosome. Table shows the position of the SNP as well as the base pair (bp) in *T. g. guttata* and *T. g. castanotis* along with the associated F_{ST} for that site.

SNP Position	Bp in *T. g. guttata*	Bp in *T. g. castanotis*	F_{ST}
chrz:32024862	T	C	0.99
chrz:32033025	T	G	0.99
chrz:32033276	T	C	0.99
chrz:32033743	T	C	0.99
chrz:32036997	A	G	0.99
chrz:32038518	A	G	0.99

Patterns of variation in song behavior

Genetic and/or cultural drift as well as selection can influence patterns of behavioral variation as it does genetic variation. It is possible that bottlenecks might reduce variation, as they often do with genetic variation (Lacy 1997; Tsutsui et al. 2000), thus, I wanted to investigate if there was a difference in the variability of song production between *T. g. castanotis* and *T. g. guttata*. For each of 17 birds that I recorded I captured between 3000 and 9000 syllables. Significant differences in 12 features were seen between the two subspecies. In general, consistent with the smaller size of the *T. g. guttata* individuals, they showed higher averages for many of the features related to frequency and pitch (Figure 2.5).

Comparison of 12 Features

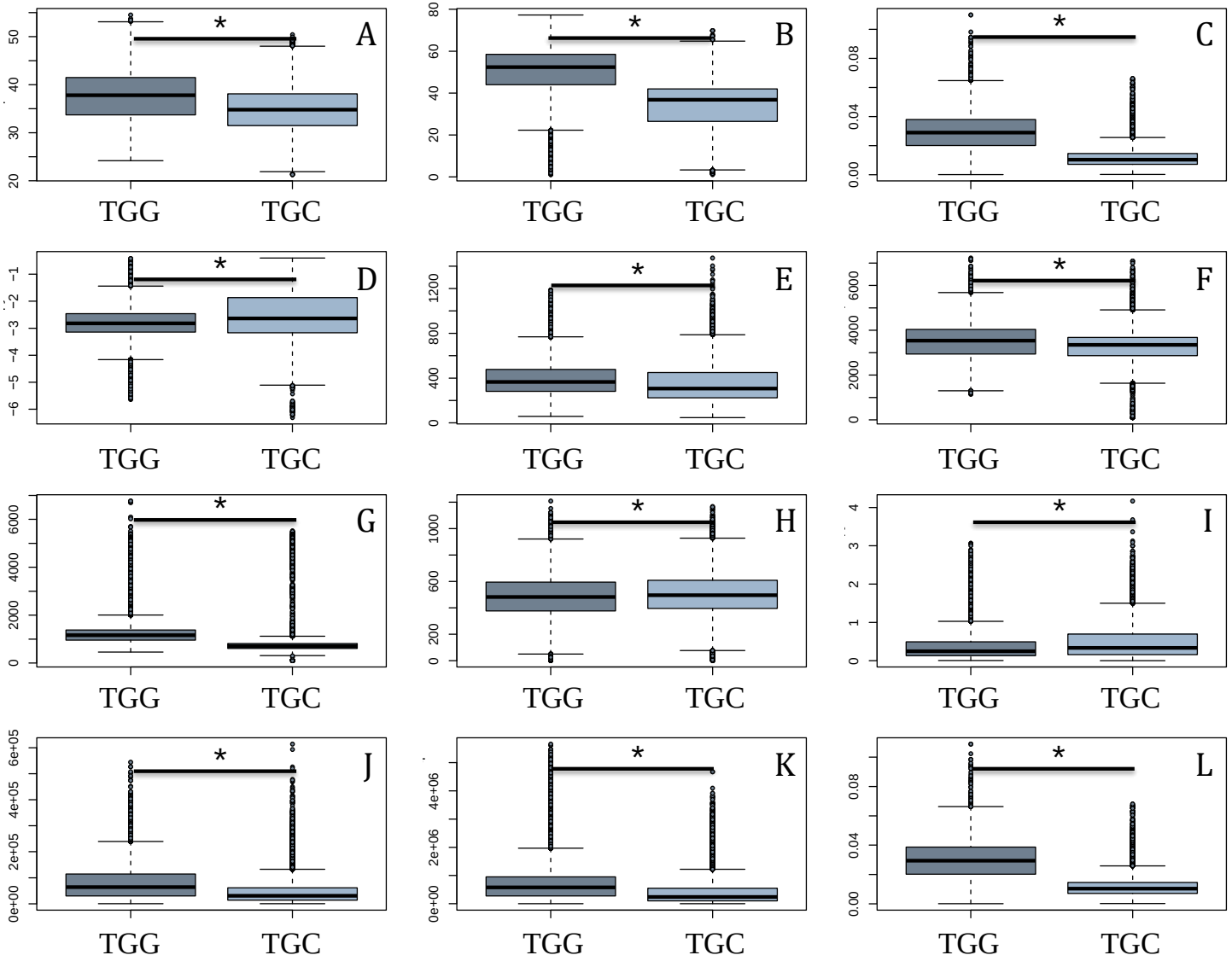

Figure 2.5. Boxplots of the 12 features measured by SAP for *T. g. guttata* and *T. g. castanotis*.

Boxplots for A) amplitude, B) mean FM, C) mean AM^2, D) mean entropy, E) mean pitch goodness, F) mean frequency, G) mean pitch, H) variance of FM, I) variance of entropy, J) variance of pitch goodness, K) variance of mean frequency, and L) variance of AM for East Carolina University *T. g. guttata* (ECU_TGG) and wild *T. g. castanotis* (Wild_TGC). Black bars with an * denote that there is a significant difference between the two populations it connects.

I also find that *T. g. castanotis* individuals are less similar to each compared with *T. g. guttata* individuals (Figure 2.6). Within the parameter plots, each cluster of dots represents a note that is repeatedly sung. Therefore, the similarity in the overall pattern of clustering between individuals visually represents how similar they are to each other. For example, for the *T. g. castanotis*, bird 1A shows a similar overall pattern to bird 1B but is less similar to 1E (Figure 2.6).

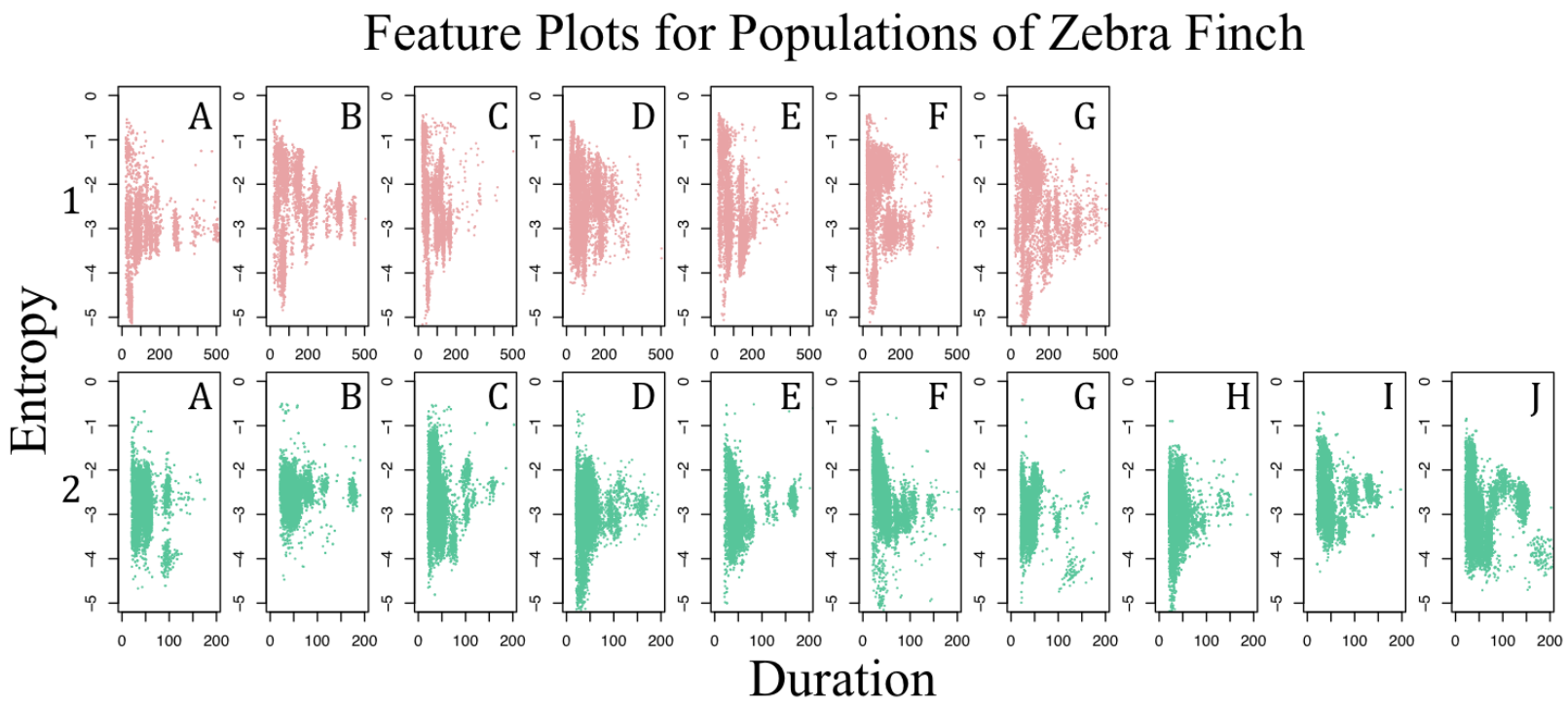

Figure 2.6. Feature plots of duration vs. entropy for the 7 *T. g. castanotis* and 10 *T. g. guttata*.

Row 1 panels A-G represent seven wild *T. g. castanotis*. And Row 2 panels A-J represent ten *T. g. guttata* from East Carolina University. Each feature plot shows duration vs. entropy. Within each parameter plot, a cluster of dots represents a note that is repeatedly sung. Therefore, the similarity in the overall pattern of clustering between each panel within a subspecies visually represents how close they are to each other.

To verify this, I ran a Mann-Whitney Wilcoxon test to see if variability of song features significantly differed between populations (if the K-L distance was significantly different). For seven out of ten comparisons, there was a statistically different K-L for the two populations in question (Table 2.2). *T. g. castanotis* showed significantly higher variability between individuals than *T. g. guttata* for all features except for amplitude, pitch, and variance of pitch goodness (Table 2.2, subset in Figure 2.7).

Table 2.2. Mean difference of K-L between *T. g. castanotis* and *T. g. guttata*.

Mean difference of K-L between populations (Mann-Whitney Wilcoxon test, p-values < 0.05 denoted with *). A positive value for mean difference signifies that the top population in the comparison has higher variability. The populations for comparison are wild *T. g. castanotis* (Wild_TGC), and East Carolina University *T. g. guttata* (ECU_TGG).

	Amp	Pitch	FM	AM^2	Ent.	Pitch Good.	Mean Freq	Var. Pitch	Var. FM	Var. Ent.	Var. Pitch Good.	Var. Mean Freq	Var. AM
Wild_TGC v ECU_TGG	**0.65**	**0.07**	**0.41**	**0.57**	**0.27**	**0.33**	**0.62**	**0.16**	**1.01**	**0.89**	**0.12**	**0.46**	**0.59**
p value	*4.47E-01*	*2.90E-01*	*9.57E-05**	*1.17E-04**	*5.28E-04**	*1.59E-02**	*1.37E-06**	*6.95E-03**	*7.76E-11**	*3.24E-09**	*1.62E-01*	*4.86E-05**	*4.37E-05**

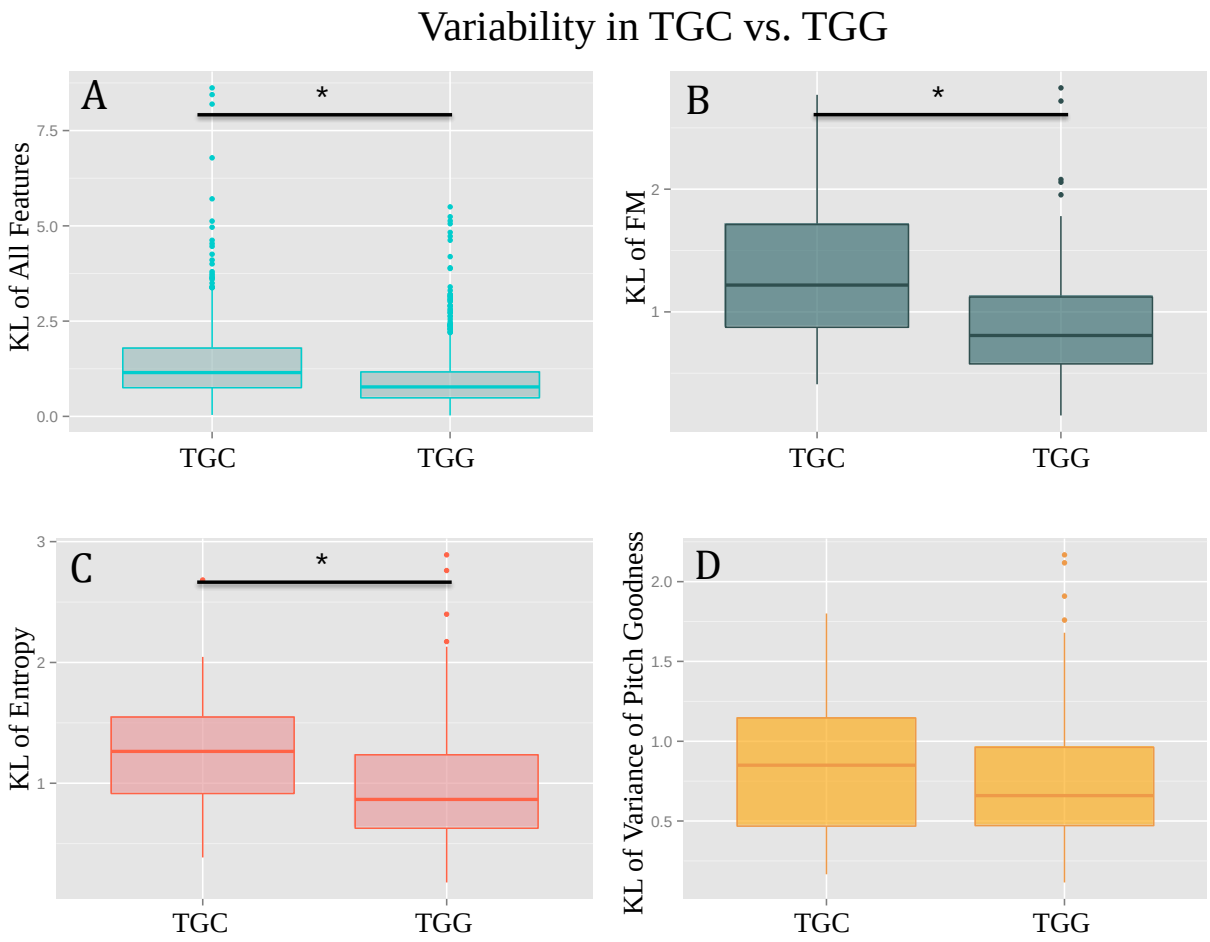

Figure 2.7. Variability (represented by K-L distance) in *T. g. castanotis* and *T. g. guttata.* K-L distance for (A) all features combined, (B) for FM, (C) Entropy, and (D) Pitch Goodness, of wild *T. g. castanotis* (TGC) and East Carolina University *T. g. guttata* (TGG). Higher K-L value denotes more variation. Black bars with an * denote that there is a significant difference between the two populations it connects.

Discussion

Despite being an important model system for neurobiology and behavior, we know little about patterns of genetic and behavioral variation among zebra finch populations. Here I quantified the level of divergence between the two subspecies of zebra finch. Using F_{ST} and diversity metrics I identified sites that vary between *T. g. castanotis* and *T. g. guttata*. I find that *T. g. guttata* shows a genome wide reduction in diversity consistent with a bottleneck event. I also conclude that the island *T. g. guttata* show a decreased variability in song between individuals compared to mainland *T. g. castanotis*. I also find a number of potential candidate SNPs that might contribute to the plumage differences between the two subspecies. Overall, my findings highlight genetic divergence is potentially due to both selection and genetic drift and point to putative functional intraspecific differences in the zebra finch. My study has helped us understand divergence of the zebra finch in allopatry.

First I wanted to investigate if the previously documented reduction of diversity in a subset of noncoding loci is present at the genome level. I found further support for a dramatic bottleneck as the two subspecies have diverged, as there is an overall reduction in diversity in the island birds. Island *T. g. guttata* show a reduction in the average *theta* for the genome. They also show a lower percentage of rare alleles than the mainland birds do. Though the overall average F_{ST} across the genome is relatively low, there are many sites that approach fixation when comparing the two subspecies. The patterns of drift would affect the two populations with varying effective population sizes very differently. I have taken the first step to investigate if this system fits under the so-called “islands of divergence” model, which posits that there will be higher divergence in areas of the genome that are under divergent selection or tightly linked to those areas (Turner et al. 2005, Nosil et al. 2009). When species are geographically isolated it is

thought that divergence occurs across the whole genome at both neutral and adaptive sites simultaneously. Subsequent to secondary contact and gene flow, the model predicts regions resistant to gene flow between species should show high differentiation and such regions tend to contain loci involved in adaptation (Turner et al. 2005). Recent studies, however, have begun to show that this simple explanation might be misleading (Noor and Bennett 2009, Renaut et al. 2012, Cruickshank and Hahn 2014) and that these islands are not driven solely by selection. Rather, genomic architecture may contribute to the generation of such islands (Noor and Bennett 2009, Renaut et al. 2012, Cruickshank and Hahn 2014). Therefore I would expect to see such islands even in allopatric species. Because local recombination tends to reduce genetic diversity, which inflates estimates of F_{ST}, I might also expect that islands of high F_{ST} occur in areas of low recombination (Cruickshank and Hahn 2014, Renaut et al. 2012, Burri et al. 2015). I have found that in the absence of homogenizing gene flow, and thus, the antagonism between selection and recombination (Kirkpatrick and Ravigné 2002, 2009, Feder et al. 2013), divergence is relatively uniform across the genome. One possibility is that these islands are simply a result of lower recombination in these areas of high F_{ST} and the next step will be to reconcile data on recombination rate (Singhal et al. 2015) with my divergence results.

In addition to a broad genome-scan, I also focused on two phenotypic traits, pigmentation and vocal behavior, in part because the pigmentation genes in particular showed multiple sites of very high F_{ST}. The island *T. g. guttata* lack the black barring along the upper breast, which is characteristic of mainland birds (Clayton et al. 1991; Zann 1996). Bills are also a darker, more intense red in *T. g. guttata* subspecies (Burley and Coopersmith 1987). Coloration in animals has many adaptive functions such as use for mate choice, competition, communication, predator avoidance, thermoregulation, and many more (Hubbard et al. 2010). Over 150 genes have been

identified that are involved in pigmentation and over 120 genes are important in humans specifically (Hubbard et al. 2010, Fernandez et al. 2008, Graf et al. 2007). I have identified a candidate SNP in exon 7 of a well-known pigmentation gene, SLC45A2. This gene is in a gene family that has gained recognition in part for its role in pigmentation is SLC, a sodium/calcium exchanger that is dependent on potassium (Hubbard et al. 2010). Two SLC genes have been identified as important in human and other vertebrate pigmentation: SLC24A5 and SLC45A2. It is known that the products from both of these genes are involved in melanin synthesis and function as transporters (Hubbard et al. 2010). Changes in the nucleotide sequences in these genes have phenotypic implications. For example, SNPs in both SLC24A5 and SLC45A2 have been found to be important in the phenotypic variation in of hair melanin (Valenzuela et al. 2010). Recently, there have been multiple studies looking at the importance of this gene in other non-model organisms. A similar non-synonymous nucleotide substitution was observed by Xu and colleagues (2013) in white tigers within SLC45A2 on exon 7, although it was at a different site with a different amino acid translation There is a functional result to this substitution: the mutation inhibits pheomelanin (Xu et al. 2013). I have also preliminarily identified six SNPs in CDKN2A, another known pigmentation gene that has been linked to barring in chickens. Sex-linked barring, which is a common plumage color in chickens, is characterized by white and black barred feathers. The white stripes are due to an absence of melanocytes (Hellström et al. 2010). Hellström and colleagues (2010) determined that the CDKN2A/B locus controlled this phenotype and CDKN2A specifically showed a near complete association. Though the potential SNPs need to be functionally investigated further, I find six sites in CDKN2A where the barred *T. g. castanotis* show a different nucleotide than the non-barred *T. g. guttata* all of which nearly at fixation between the two populations.

The second phenotypic I focused on in terms of intraspecific variation was song production. As was found by Clayton (1990c), I see differences in the songs produced by the two subspecies. She found differences in the length of the song, the frequency, and the number of elements in a phrase. The *T. g. guttata* sang songs that lasted 0.95 seconds whereas *T. g. castanotis* songs were 0.3 seconds shorter. *T. g. guttata* songs were a slightly higher frequency, and included phrases with 11 elements instead of the 8 elements seen in *T. g. castanotis*. Hybrid songs were intermediate for these features between the two subspecies (Clayton 1990c). In my study I saw that for the average notes of *T. g. guttata* songs they showed increased amplitude, higher mean frequency modulation, higher mean AM^2, higher mean pitch goodness, higher mean frequency, higher mean pitch, bigger variance of pitch goodness, higher variance of mean frequency, and higher variance of AM. The higher frequencies are likely due to the smaller body size and are consistent with the Zann (1983) study that found that the distance calls in *T. g. guttata* have a higher fundamental frequency. I also found a difference in the variability of songs among individuals between the two subspecies: the island subspecies shows less variation in song among individuals than the mainland birds. It is possible that this reduction in variability is due to the founder effect that led to only a subset of the variety being represented in the island populations as well as probable subsequent cultural drift. Additionally, the higher variation in *T. g. castanotis* could be due to selective pressures as well. Zebra finches occur in huge numbers in Australia (Zann 1996) so potentially there has been selection on variability for individual recognition, which might not occur as strongly in *T. g. guttata*. Zann (1993) described variation of 33 wild Australian populations both between and within geographic zones. He found that the song structure varied among populations at a macrogeographic level but not a microgeographic

level, which he attributed to dispersal between colonies (Zann 1993). In terms of selection, it could also be that *T. g. guttata* show reduced sexual selection, as is true with many other island species. There is evidence of the weaker role sexual selection plays on islands seen in lower frequency of extra pair paternity in socially monogamous passerines (Price 2008). If sexual selection favors song diversity, and is reduced on islands, that could explain the reduction in variability in *T. g. guttata* songs. Overall, my findings add support for the occurrence of a severe bottleneck affecting both the genome and aspects of behavior when the Lesser Sunda Islands were colonized and highlight genetic divergence is potentially due to both selection and genetic drift. In general, my study has taken an important first step to characterize the intraspecific genetic and behavioral variation in this system and has begun to describe divergence in the context of allopatry with no secondary contact.

References

Balakrishnan CN, Edwards SV (2009) Nucleotide variation, linkage disequilibrium and founder-facilitated speciation in wild populations of the zebra finch (Taeniopygia guttata). *Genetics*, 181:645-60.

Burley N, Coopersmith CB (1987) Bill Color Preferences of Zebra Finches. *Ethology*, 76(2):133–151.

Burrell AS, Disotell TR, Bergey CM (2015) The use of museum specimens with high-throughput DNA sequencers. *Journal of human evolution*, *79*:35-44.

Burri R, Nater A, Kawakami T, Mugal CF, Olason PI, Smeds L, Suh A, Dutoit L, Bureš S, Garamszegi LZ, Hogner S (2015) Linked selection and recombination rate variation drive the evolution of the genomic landscape of differentiation across the speciation continuum of Ficedula flycatchers. *Genome research*, *25*(11):1656-1665.

Catchpole CK, Komdeur J (1993) The song of the Seychelles warbler Acrocephalus sechellensis, an island endemic. *Ibis*, *135*(2):190-195.

Clayton NS (1989) The effects of cross-fostering on selective song learning in estrildid finches. *Behaviour*, 190:163-74.

Clayton NS (1990a) Assortative Mating in Zebra Finch Subspecies, Taeniopygia guttata guttata and *T. g. castanotis*. *Philosophical Transactions: Biological Sciences*, 330(1258):351-370.

Clayton NS (1990b) Mate choice and pair formation in Timor and Australian Mainland zebra finches. *Anim. Behav*., 39: 474~480.

Clayton NS (1990c) Subspecies recognition and song learning in zebra finches. *Anim. Behav*., 40:1009-1017.

Clayton DF (2004) Songbird genomics: methods, mechanisms, opportunities, and pitfalls. *Annals of the New York Academy of Sciences*, 1016(1):45-60.

Clayton DF, Huecas M (1990) Forebrain-enriched RNAs of the canary: a population analysis using hybridization kinetics. *Molecular Brain Research*, 7:23-30.

Clayton NS, Hodson D, Zann R (1991) Geographic variation in zebra finch subspecies. *Emu*, 91: 2–11.

Coyne JA, Orr HA (2004). *Speciation*. Sunderland, MA.

Cruickshank TE, Hawn MW (2014) Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Molecular Ecology*, 23:3133-57.

Dave AS, Margoliash D (2000) Song Replay During Sleep and Computational Rules for Sensorimotor Vocal Learning. *Science*, 290(5492):812-816.

Feder JL, Flaxman SM, Egan SP, Comeault AA, Nosil P (2013) Geographic Mode of Speciation and Genomic Divergence. *Annual Review of Ecology, Evolution, and Systematics*, 44:73-97.

Fernandez LP, Milne RL, Pita G, et al. SLC45A2: a novel malignant melanoma-associated gene. *Hum Mutat*., 29(9):1161–7.

Forstmeier W, Segelbacher G, Mueller JC, Kempenaers B (2007) Genetic variation and differentiation in captive and wild zebra finches (*Taeniopygia guttata*). *Molecular ecology* 16:4039–50.

Forstmeier W, Burger C, Temnow K, Derégnaucourt S (2009) The genetic basis of zebra finch vocalizations. *Evolution* 63, 2114-2130.

Gilby AJ, Mainwaring MC, Griffith SC (2013) Incubation behaviour and hatching synchrony differ in wild and captive populations of the zebra finch. *Anim Behav*, *85*(6):1329-1334.

Graf J, Voisey ÃJ, Hughes I, Daal A Van (2007) Graf, J., Voisey, J., Hughes, I., & Van Daal, A. (2007) Promoter polymorphisms in the MATP (SLC45A2) gene are associated with normal human skin color variation. *Human mutation*, 28(7):710-717.

Griffith SC (2000) High fidelity on islands: a comparative study of extrapair paternity in passerine birds. *Behavioral Ecology*, *11*(3):265-273.

Hellström AR, Sundström E, Gunnarsson U, Bed'Hom B, Tixier- Boichard M, Honaker CF, ... & Kerje S (2010) Sex- linked barring in chickens is controlled by the CDKN2A/B tumour suppressor locus. *Pigment cell & melanoma research*, *23*(4):521-530.

Hoskin CJ, Higgie M, McDonald KR, Moritz C (2005) Reinforcement drives rapid allopatric speciation. *Nature*, *437*(7063):1353-1356.

Hubbard JK, Uy JAC, Hauber ME, Hoekstra HE, Safran RJ (2010) Vertebrate pigmentation : from underlying genes to adaptive function. *Trends in Genetics*, *26*(5):231–239.

Hung CM, Shaner PJL, Zink RM, Liu WC, Chu TC, Huang WS, Li SH (2014) Drastic population fluctuations explain the rapid extinction of the passenger pigeon. *Proceedings of the National Academy of Sciences*, *111*(29):10636-10641.

Jin H, Clayton DF (1997) Localized Changes in Immediate-Early Gene Regulation during Sensory and Motor Learning in Zebra Finches. *Neuron*, 19:1049–1059.

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Mentjies P, Drummond A (2012) Geneious Basic:

an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics, 28*(12):1647-1649.

Kirkpatrick M, Ravigné V (2015) Speciation by Natural and Sexual Selection: Models and Experiments. *The American Naturalist,* 159:S22-S35.

Korneliussen TS, Albrechtsen A, Nielsen R (2014) ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinformatics,* 15:1-13.

Kuehne HA, Murphy HA, Francis CA, Sniegowski PD (2007) Allopatric divergence, secondary contact, and genetic isolation in wild yeast populations. *Current Biology, 17*(5):407-411.

Lacy RC (1997) Importance of genetic variation to the viability of mammalian populations. *Journal of Mammalogy,* 78(2):320-335.

London SE, Clayton DF (2008) Functional identification of sensory mechanisms required for developmental song learning. *Nature Neuroscience,* 11:579-586.

Marler P (1970) Birdsong and speech development: could there be parallels? *American Scientist,* 58:669-73.

Mayr E (1944) Timor and the colonization of Australia by birds. *Emu,* 44:113-30.

McCormack JE, Tsai WL, Faircloth BC (2016) Sequence capture of ultraconserved elements from bird museum specimens. *Molecular ecology resources, 16*(5):1189-1203.

Grant PR, Grant BR (2009) The secondary contact phase of allopatric speciation in Darwin's finches. *Proceedings of the National Academy of Sciences, 106*(48):20141-20148.

Mooney R (2009) Neural mechanisms for learned birdsong. *Learning and Memory,*16:655–69.

Nosil P**,** Funk DJ, and Ortiz-Barrientos D (2009) Divergent selection and heterogeneous genomic divergence. *Molecular Ecology,* 18:375-402.

Noor MAF, Bennett SM (2009) Islands of Speciation or Mirages in the Desert? Examining the Role of Restricted Recombination in Maintaining Species. *Heredity,* 103(6):439-444.

Olveczky BP, Andalman AS, Fee MS (2005) Vocal experimentation in the juvenile songbird requires a basal ganglia circuit. *PLoS Biology,* 3:e153.

Price T (2008) *Speciation in birds*. Roberts and Company Publishers.

Renaut S, Grassa CJ, Yeaman S, Moyers BT, Lai Z, Kane NC, Bowers JE, Burke JM, Rieseberg LH (2012) Genomic islands of divergence are not affected by geography of speciation in sunflowers. *Nature Communications,* 4:1-8.

Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP (2011) Integrative Genomics Viewer. *Nature Biotechnology*, 29:24–26.

Slater PJB, Eales LA, Clayton NS (1988) Song learning in zebra finches (Taeniopygia guttata): progress and prospects. *Advances in the Study of Behavior*, 18:1-34.

Sound Analysis Pro http://ofer.sci.ccny.cuny.edu/html/sound_analysis.html

Tchernichovski O, Nottebohm F, Ho CE, Pesara B, and Mitra PP (2000) A procedure for an automated measurement of song similarity. *Animal Behaviour*, 59:1167-76.

ten Cate C (2014) On the phonetic and syntactic processing abilities of birds: From songs to speech and artificial grammars. *Current Opinion in Neurobiology*, 28:157–164.

Thompson JA, Basista MJ, Wu W, Bertram R, Johnson F (2011) Dual Pre-Motor Contribution to Songbird Syllable Variation. *The Journal of Neuroscience*, 31(1):322–330.

Thorvaldsdóttir H, Robinson JT, Mesirov JP (2013) Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in Bioinformatics*, 14:178-192.

Tsutsui ND, Suarez AV, Holway DA, Case TJ (2000) Reduced genetic variation and the success of an invasive species. *Proceedings of the National Academy of Sciences*, *97*(11):5948-5953.

Turner TL, Hahn MW, Nuzhdin S V (2005) Genomic islands of speciation in Anopheles gambiae. *PLoS biology*, 3:e285.

Valenzuela RK, Henderson MS, Walsh MH, Garrison NA, Kelch JT, Cohen-Barak O., … Brilliant MH (2010) Predicting phenotype from genotype: normal pigmentation. *Journal of Forensic Sciences*, *55*(2):315–22.

Verner J (1964) Evolution of polygamy in the long-billed marsh wren. *Evolution*, *18*(2):252-261.

Via S (2009) Natural selection in action during speciation. *Proceedings of the National Academy of Sciences*, *106*(Supplement 1):9939-9946.

Warren WC, Clayton DF, Ellegren H, Arnold AP, Hillier LW, Künstner A, ... & Heger A (2010) The genome of a songbird. *Nature*, *464*(7289):757-762.

Xu X, Dong GX, Hu XS, Miao L, Zhang XL, Zhang DL, Yang HD, Zhang TY, Zou ZT, Zhang TT, Zhuang Y (2013) The genetic basis of white tigers. *Current biology*, 23(11):1031-5.

Zann RA (1993) Variation in song structure within and among populations of Australian zebra finches. *The Auk*, 110(4):716-726.

Zann RA (1996) *The zebra finch: a synthesis of field and laboratory studies*. Oxford: Oxford University Press.

"Zebra Finch - *Taeniopygia guttata castanotis*". http://www.efinch.com/species/zebra.html.

CHAPTER 3 A COMPARISON OF SONG LEARNING IN THE TWO ZEBRA FINCH SUBSPECIES (*TAENIOPYGIA GUTTATA*)

Abstract

Learned vocal communication, or vocal learning, is a trait that is shared by humans and songbirds, but rare in other animals. Unlike innate communication, learned vocalizations are acquired early in life by juveniles listening and copying what they hear from adults. Zebra finches have served as the dominant model for vocal learning for over half a century. There are two subspecies of zebra finches that differ in song variability, with the domesticated *Taeniopygia guttata castanotis* showing more variability among individuals compared to both the other subspecies *Taeniopygia guttata guttata* and the wild *T. g. castanotis*. In order to differentiate between genetic or cultural controls of this difference in variation, I cross-fostered both subspecies to the Bengalese finch, *Lonchura striata domestica*, to test for differences in song copying behavior. I found that we cannot reject the null hypothesis that zebra finches copy equally, although it appears that the variability among the *T. g. castanotis* individuals remains greater whether or not they are cross-fostered to a different species.

Introduction

Introduction to Vocal Learning & Study System

Vocal learning in early development is a trait that is shared by humans and songbirds (Mello 2014), but rare in other animals (Doupe and Kuhl 1999). Songbirds include about 4000 species, though song learning has been studied in less than 10% of these species (Mooney 2009).

For over half of a century, the zebra finch (*Taeniopygia guttata*) has been the dominant model for vocal learning (Slater et al. 1988; Forstmeier et al. 2009) and an important model for understanding the origins of the complexity of the human language (Marler 1970; ten Cate 2014). In songbirds such as the zebra finch, the male learns to sing from its tutor (usually its parent, or in manipulative experiences, a foster parent) during a critical period. In natural situations song learning is constrained by environmental stress (Buchanan et al. 2004) and it is controlled by degrees of brain development (Nowicki et al. 2002). Song learning occurs in two main phases: sensory learning and sensorimotor learning. During the sensory learning stage the nestling listens and memorizes the tutor's song (days ~ 20-60). Next, during the sensorimotor period, which overlaps the sensory learning stage (days ~ 32-90), the young bird practices its song and will attempt to bring its song closer to the tutor's with each iteration. It is not critical that the young bird can hear the tutor's song during this period, meaning that it can be accomplished through memory in some cases (Mooney 2009).

While much is known about the neural mechanisms for learned vocal communication (Jin and Clayton 1997; Dave and Margoliash 2000; Olveczky et al. 2005; Mooney 2009; London and Clayton 2008; Thompson et al. 2011; Vallentin et al. 2016), little attention has been paid to the mechanisms that contribute to variation in song learning or production. Therefore, although zebra finch is a classic model system, work on these birds has mostly overlooked potential genetic variation that exists in the traits that are being studied. Indeed, most research to date has centered exclusively around one of the two subspecies, the Australian zebra finch, *Taeniopygia guttata castanotis*, using domesticated birds that have been bred in captivity for around 150 years. The second subspecies is the Timor zebra finch, *Taeniopygia guttata guttata*, which was brought into captivity only about 20 years ago (Rogers 1979). *T. g. castanotis* is found on the

mainland of Australia whereas the *T. g. guttata* subspecies is found on the islands north of the continent (referred to as the Timor or Lesser Sunda Islands) (Mayr 1944; Clayton 1990a-c). It is estimated that *T. g. guttata* colonized the lesser Sunda islands 1-2 MYA (Balakrishnan and Edwards 2009, Newhouse and Balakrishnan 2015).

Song Variation in Zebra Finches

Previous work has sought to characterize the patterns of variation in song behavior between these two populations. *T. g. guttata* have significantly less variation among individuals than both wild and domesticated *T. g. castanotis* (chapter 1 and 2). Although zebra finches prefer to mate assortatively by subspecies and the females prefer the songs of their own subspecies (Clayton 1990a; Clayton1990b) as well as prefer to associate with and respond to conspecifics (Campbell et al. 2009), they will interbreed in captivity and produce viable, fertile offspring (Davidson and Balakrishnan 2016). Clayton (1990c), found that there were differences between the songs of the two subspecies. She found differences in the length of the song, the frequency, and the number of elements in a phrase. Differences in frequency were attributed to differences in body size seen between the subspecies, as *T. g. guttata* is significantly smaller. The *T. g. guttata* sang songs that lasted 0.95 seconds, had a slightly higher frequency, and included phrases with 11 elements (Clayton 1990c). The distance calls, which are produced by males and females and are shorter and simpler than songs and used for communication of threats or location, in *T. g. guttata* also have a higher fundamental frequency (the lowest frequency in the oscillation) (Zann 1983). In contrast, *T. g. castanotis* songs were on average 0.3 seconds shorter, slightly lower in frequency, and had phrases with eight elements. Hybrid songs were intermediate between the two subspecies (Clayton 1990c). Clayton (1990c) also conducted

inverse cross-fostering experiments between the two subspecies (cross-fostering each subspecies to the other subspecies), and found that the same differences in song held for both of the subspecies whether they were raised by conspecifics or heterospecifics, suggesting that rearing experience had little effect on macrostructural differences in song. This led her to hypothesize that those differences are used for subspecies recognition (Clayton 1990c). My study builds on this work by investigating the genetic underpinning of differences in song between the subspecies

Cognitive Ability and Song

Birdsong is a sexually selected trait (Hawkins 1918) that is involved in courtship, maintaining pair bonds, and territory defense (Suzuki et al. 2014). Females often prefer males that sing a more complex song (Kroodsma 1976; Catchpole 1980, 1986, 1996; Hiebert et al. 1989; Searcy 1992; Lampe and Saetre 1995; Gentner and Hulse 2000; Okanoya 2004a,b; Leitão et al., 2006). The costs of complex songs, expressed as physical and developmental constraints, are related to energy and time consumption, aggression, and other costs (Gil and Gahr 2002). More neural capacity is necessary in the brain to store more song elements (Honda and Okanoya 1999; Nottebohm et al. 1981), so complexity serves as an honest signal of male quality (Zahavi 1975; Grafen 1990; Johnstone and Grafen 1993; Buchanan et al. 2004, Peters et al. 2014). Additionally, in the wild, singing loud, complex songs could potentially be costly due to making the individual more vulnerable to predators, drawing on cognitive resources which are also involved in reacting to dangers, and the cost of evolving and maintaining brain mechanisms that underlie complex songs (Okanoya 2012). Peters and colleagues (2014) found that there were parallel impacts of developmental stress on song learning and cognition, which lends credence to

the associations of cognitive ability and song. This supports the hypothesis that cognitive ability is indicated by the sexually selected qualities of song (Peters et al. 2014). There is a positive correlation between the complexity of song and chick weight, parental effort by the male (Buchanan and Catchpole 2000) and offspring survival (Hasselqvist et al. 1996). Therefore, females are selecting males with complex songs to benefit their offspring, and complexity may evolve in part due to sexual selection through female choice (Anderson and Iwasa 1996; Okanoya 2002). For example, wild male song sparrows *Melospiza melodia* with larger repertoires had higher fitness (Reid et al. 2004) and a similar result was found in sedge warblers, *Acrocephalus schoenobaenus* (Buchanan and Catchpole (1997). In captivity, male zebra finches with more elements per song were found to require fewer learning trials to perform a novel foraging task (Boogert et al. 2008). However, in a recent meta-analysis Soma and Garamszegi (2011) found a weak and independent mean effect size for the song/mating success association and strength of sexual selection. Therefore, Soma (2011) cautions against assuming that the reproductive advantage of males with complex songs is prevalent in most Oscines (songbirds).

Genetic and Learned Components of Song

Zebra finches and other songbirds have learned as well as innate components of their vocalizations (Mooney 2009). Female calls, which are not learned, are highly heritable (Forstmeier et al. 2009). Male calls and songs, however, show lower heritability, as learning adds variability (Forstmeier et al. 2009; Woodgate et al. 2013). In a remarkable study, Fehér et al. (2009) showed that zebra finch colonies founded from birds with experimentally tutored "isolate" songs, returned to wild-type song culture in only a few generations, suggesting a genetic basis for song culture. Mori and Wada (2015) found a similar result where the songs of

deafened birds eventually still crystalized revealing that some of vocal development and stabilization in zebra finch was independent from sensory inputs and instead due to developmental genetic programs. In addition to the structure of the song itself, aspects of underlying neural substrates (i.e. song nuclei) have also been documented to be heritable (Airey et al. 2000). Voice characteristics such as frequency and timbre function as honest indicators of the size of the bird; in contrast, structural traits such as repertoire size in males showed low heritability (Forstmeier et al. 2009). Culture, including song culture (members of a species showing variation in song, and geographically tractable dialects) is typically associated with inheritance through social learning. However, species-specific genetic constraints are involved as well (Marler and Tamura 1962; Nelson 2000; Fehér et al. 2009).

Many songbirds have an innate bias towards learning and producing conspecific song even when tutored artificially using another species' song (Marler and Peters 1977; Podos et al. 2004). Bengalese-fostered zebra finches learn few notes and have shorter bout lengths with fewer repeated notes than Bengalese finches. Element morphology was incomplete in the fostered zebra finches, but they learned element structure well (Takahasi et al. 2006). This shows that acquisition of song is accomplished through a combination of innate influences, learning from the social context, and slight individual variation of the song when compared to the tutor's song (Clayton 1989).

Thus it is clear that both genes and environment contribute to vocal behavior. The vast majority of song learning research has focused on the mechanisms of learning, but no studies have examined population variation in song copying ability in the zebra finch. I seek to resolve alternative hypotheses for differences in song variability in captive colonies of *T. g. guttata* and *T. g. castanotis* and to determine why there is increased variability in *T. g castanotis* in general,

and an even higher variability in the domesticated *T. g. castanotis* songs (see chapter 1 and 2). One possibility is that this trend is due to a cultural shift. Alternatively, it is possible that a genetic shift has occurred leading to the change in variability. In order to distinguish between the possible explanations of increased variation among domestic *T. g. castanotis* individuals, I conducted a cross-fostering experiment to test song copying ability in the two subspecies. If both subspecies show similar song copying ability that would suggest that the observed differences in song variability can be explained by a cultural (behavioral) shift rather than a genetic shift (selection or drift). In order to compare the differences in the subspecies' ability to learn songs, I tested to see how well the two subspecies of zebra finch copy the song of a third species, the Bengalese finch (*Lonchura striata domestica*).

First, I test if the songs of cross-fostered birds are more similar to those of their tutors' or songs of conspecifics. This is in order to estimate the relative degree of learning that has taken place for each subspecies. Secondly, I will test if time of cross-fostering, specifically cross-fostering at the egg stage vs. fledgling stage at approximately day 30 after hatch (P30), led to different results as to how similar songs were to tutors. This will test if the amount of time in the sensitive phase that the bird is exposed to the Bengalese tutor affects the degree to which the tutor's song is learned for each subspecies. Third, I compare the cross-fostered songs to the tutor songs (breaking down the comparison by time of cross-fostering (all vs. egg- vs. fledgling-cross-fostered). This test allows me to see which features of song are more similar to the tutor's and which subspecies has more accurately learned each feature. Then I compare variability of song among the tutors; this is to test if the Bengalese tutors vary significantly from each other in terms of their songs. Subsequently, I compared cross-fostered birds to other non-tutor Bengalese finch songs in order to see if cross-fostered birds more closely matched their specific tutor to further

test for learning. Lastly, I ran a within-subspecies comparison of the cross-fostered birds to each other. This was to see if the result that *T. g. castanotis* songs are more individually distinctive held true despite cross-fostering.

Methods

Study Populations

The East Carolina University populations of *T. g. guttata* and domesticated *T. g. castanotis* birds were established in 2012. Originally, the domesticated *T. g. castanotis* colony was founded from five pairs of birds derived from another captive research colony at ECU. The *T. g. guttata* colony was originally founded at the University of Illinois with five pairs of birds. The descendants of this founding colony (~30 birds) were moved to ECU in 2012. Roy Beckham brought the *T. g. guttata* into captivity 25 years ago (efinch.com).

The two subspecies were each housed in large flight aviaries separated by subspecies at the Brody School of Medicine. Birds had *ad libitum* access to commercial finch dry-seed mix, with fresh water provided daily and finely crush warmed hard-boiled eggs and greens given weekly. The birds were monitored on a daily basis. Housing and experimental protocols were approved by the East Carolina University Institutional Animal Care and Use Committee. I used songs from five birds from each subspecies recorded for chapter 1 and 2 to serve as non-cross-fostered control groups.

Experimental Cross Fostering

In order to compare song-copying behavior between two subspecies, I conducted a cross-fostering experiment where zebra finches were tutored by Bengalese finches, *Lonchura striata*

domestica. This approach has been used successfully used many times in experiments with zebra finches (e.g., Eales 1987a; ten Cate 1987; Clayton 1989; Takahasi et al. 2006; Campbell and Hauber 2008; Campbell and Hauber 2009; Soma 2011; Villian et al. 2015).

Throughout this study I used five Bengalese Finch males as tutors. I used two different strategies for cross-fostering, moving eggs from parents to foster parents (n=5), and moving birds at ~ 30 days post hatch to tutoring cages (n=13). I employed primarily this latter approach because I was having poor success with egg swapping (many clutches were being abandoned). For those individuals who were successfully reared from the egg stage, I left the *T. g. guttata* or *T. g. castanotis* birds in with the Bengalese tutor until post-hatch day 90.

For the latter approach, birds that were allowed to hatch under the care of their genetic parents were genetically sexed using standard protocols with the P2-P1237L primers (Griffiths et al. 1998; Khan et al. 1998; Jensen et al. 2003; Ong and Vellayan 2008; Kolts and McRae 2017). Then, once the *T. g. castanotis* and *T. g. guttata* male fledglings reached this age I moved one male in each species into a cage with a Bengalese male and left them there until day 90 when the zebra finch song crystalizes. One *T. g. guttata* fledgling separated from his tutor around day 50 after showing extreme aggression towards the Bengalese male. He was quartered off into a section of the cage where he could still see and hear the tutor. A female Bengalese was placed with the male Bengalese to help with stress.

Song Recordings

Once the chicks reached full sexual maturity, the songs of the Bengalese tutors as well as the male foster chicks were recorded in a sound chamber. A pair of birds, one male and one female, was placed in the sound chamber and left overnight to be recorded using Sound Analysis

Pro (SAP) software (Tchernichovski et al. 2000, http://ofer.sci.ccny.cuny.edu/html/sound_analysis.html). The study of song learning has relied heavily on the development of software to analyze recordings and produce sonograms to look at similarities and differences between songs (Mooney 2009). Tchernichovski and colleagues were able to investigate how a young bird's song differed note-to-note from the tutor's song, pioneering this type of study (Tchernichovski et al. 2001). SAP allows for activity-triggered recordings, making it easier to select out recording with vocalizations. The chamber was opened once a day to check food and water and the pair was left in the chamber for an average of three days or until the males produced at least 100 song recordings. I found that the *T. g. guttata* subspecies do not sing well when isolated so I used a male and female instead which means that I captured both directed (to the female) and undirected song (not specifically directed to the female). Undirected song is slightly more diverse within-individual than directed, but Woolley and Doupe (2008) suggest that the within-individual differences are very subtle. Therefore, following previous work that has grouped both type of song together (Lachlan et al. 2016) I do not expect this to influence my overall results. Once the recordings were collected, I manually sorted through the files to select out only those that contained songs and discarded other vocalizations or noises from movement.

Sounds Analysis Pro was also used for quantitative comparisons of song structure. As a first measure of song similarity, I used the Kullback-Leibler divergence in pairwise comparisons among individuals (Wu et al. 2008). A higher K-L distance indicates more spread in the data and therefore songs that are more dissimilar among the individuals being compared. In cases where there are high levels of stereotypy, this can indicate that there is a high success in learning (Deregnaucourt et al. 2005). To accomplish this, the Feature Batch function in SAP was first used to parse the motifs into syllables by setting certain segmentation values, which are unique to

each individual. To do this, I opened a subset of ~20 songs in SAP per individual and adjusted the segmentation setting for one or two features until each syllable was more effectively separated out. The features that can be used to segment are amplitude, pitch, mean frequency, goodness of pitch, FM, AM, Wiener entropy, continuity (t), and continuity (f). In most cases amplitude, sometimes with a secondary feature of mean frequency or continuity, was used to segment the song recordings into distinct syllables. Once the settings were selected for segmentation, SAP automatically segmented all the song recordings for that individual bird into separate syllables for further comparison.

All K-L divergence estimates include syllable duration as one of the variables, but were estimated for a suite of different secondary song parameters. The batch analysis from SAP creates syllable tables which give information for 14 parameters: duration, mean amplitude, mean pitch, mean FM, mean AM^2, mean entropy, mean pitch goodness, mean frequency, variance in pitch, variance in FM, variance in entropy, variance in pitch goodness, variance in mean frequency, and variance in AM. I filtered out the feature data for syllables that were under 0.2 seconds to account for segmentation errors that resulted in only a fraction of a syllable. I took five main approaches to examine K-L divergence among populations. First, for each cross-fostered individual I calculated K-L divergence directly based between the individual and its tutor. Second, I also computed a pairwise K-L divergence matrix among tutored individuals from five individuals within each subspecies. This latter approach provides a measurement of how different the tutored birds are from each other, rather than from the tutor. For this second comparison I used a subset of birds that had a matched set of tutors. Third, I compared K-L between the cross-fostered birds to their tutors and K-L of the cross-fostered birds to conspecifics (cross-fostered *T. g. castanotis* to *T. g. castanotis* and cross-fostered *T. g. guttata* to

T. g. guttata). When comparing the zebra finches to their tutors I compared all cross-fostered *T. g. castanotis* males vs. all cross-fostered *T. g. guttata*, egg cross-fostered *T. g. castanotis* vs. egg cross-fostered *T. g. guttata*, and fledgling cross-fostered *T. g. castanotis* vs. fledgling cross-fostered *T. g. guttata*. Forth, I compared five cross-fostered zebra finches of each sub-species to a randomly assigned tutor to see if they were more similar to their Bengalese tutor than a non-tutor Bengalese finch. Finally, I also compared the cross-fostered zebra finch to each other within subspecies.

Statistical Analyses

All statistical analyses were performed using the software R (R Core Team 2014). Statistical analyses were performed on approximately 100 songs per 23 individuals, (18 cross-fostered zebra finches and five Bengalese finches). After calculating the K-L distance for the 13 song features, I ran a Shapiro-Wilk Normality Test on the data. Since most of the data were not normally distributed, I used a Mann Whitney Wilcoxon test to determine if the K-L distances for each feature were significantly different for the comparisons.

Results

Similarity of Cross Fostered Birds to Their Tutors

In total, I was able to successfully cross-foster ten male *T. g. castanotis* and eight male *T. g. guttata* to maturity and record their songs. For each bird I selected out 100 song recordings, with between 3000 and 9000 syllables.

In order to investigate the effect of cross-fostering, I compared the cross-fostered birds to their tutors, as well as to conspecifics from the general (non cross-fostered) *T. g. castanotis* and *T. g. guttata* ECU populations. I used five individuals each of non-cross-fostered *T. g. castanotis* and *T. g. guttata* (chapter 1 and 2) as the templates from which to calculate K-L for the cross-fostered *T. g. castanotis* and *T. g. guttata* respectively.

For two features (AM^2 and variance of AM), cross-fostered *T. g. castanotis* were significantly more similar to their tutors than they were to normally raised *T. g. castanotis*. Two other features (entropy and variance of pitch goodness) showed the opposite pattern where the cross-fostered *T. g. castanotis* were more similar to normally raised *T. g. castanotis* than the tutors that raised them (Table 3.1, Figure 3.1A).

For seven features (pitch, FM, AM^2, entropy, pitch goodness, variance of entropy, and variance of AM), cross-fostered *T. g. guttata* were more similar to normally raised *T. g. guttata* than they were to their tutors (Table 3.1, Figure 3.1B). For these seven features, the K-L distance was higher between cross-fostered birds and their tutors, meaning that they were more similar to their own species even though Bengalese males raised them.

Table 3.1. Mean K-L for *T. g. guttata* and *T. g. castanotis* compared to their tutors and to conspecifics.

Mean K-L for the 13 features. Comparison of the cross-fostered birds to tutors vs. cross-fostered birds to normally raised conspecifics (Mann-Whitney Wilcoxon test, p-values < 0.05 denoted with *). Above the double line are the *T. g. guttata* birds and below are the *T. g. castanotis* birds. Higher K-L means greater variability.

	Amp	Pitch	FM	AM^2	Ent.	Pitch Good.	Mean Freq	Var. Pitch	Var. FM	Var. Ent.	Var. Pitch Good.	Var. Mean Freq	Var. AM
XF_TGG to Tutor (Mean)	3.49	3.25	3.02	2.86	3.83	3.65	2.29	0.33	1.38	1.99	0.97	1.53	2.77
XF_TGG to TGG (Mean)	2.29	1.75	1.25	1.25	1.28	1.14	1.32	0.43	0.74	0.98	0.85	1.04	1.20
p-value	0.09	0.02 *	0.01 *	0.001 *	2.12e-08*	2.13e-05*	0.44	0.08	0.22	0.004 *	0.65	0.16	0.001 *
XF_TGC to Tutor (Mean)	4.45	3.84	2.61	2.14	4.16	3.5	2.85	0.72	1.97	2.35	1.42	1.46	2.14
XF_TGC to TGC (Mean)	5.9	2.89	2.82	3.16	2.51	3.001	2.54	0.69	1.94	2.44	2.16	2.08	3.17
p-value	0.22	0.17	0.82	0.04*	0.008 *	0.59	0.76	0.8	0.47	0.14	0.01*	0.06	0.02*

XF to Conspecifics and Tutors

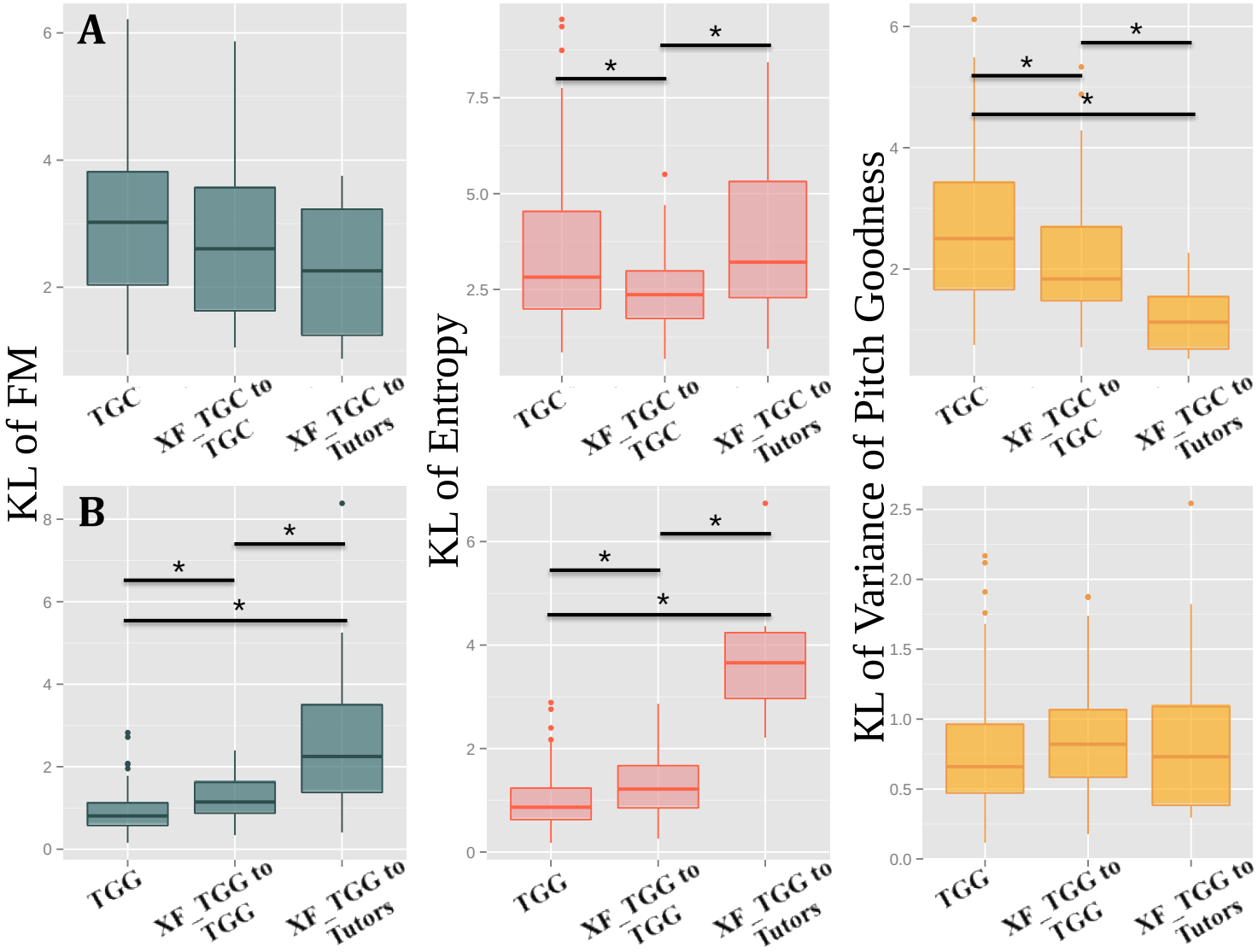

Figure 3.1. K-L for three features for within un-fostered zebra finches, cross-fostered zebra finch to conspecific, and cross-fostered zebra finch to their tutors.

K-L for FM, entropy, and variance of pitch goodness for within un-fostered zebra finches, cross-fostered zebra finch to conspecific, and cross-fostered zebra finch to tutor for a) *T. g. castanotis* (TGC) and b) *T. g. guttata* (TGG). For each subspecies, two of the highlighted features show a significantly different K-L between the comparisons. Black bars with an * denote that there is a significant difference between the two populations it connects.

T. g. guttata had more significant features in common with normally raised conspecifics than *T. g. castanotis* did. This pattern could mean that more features are innate in *T. g. guttata* and therefore not affected as strongly by cross-fostering.

Cross fostering from Eggs Versus Fledgling stage P30

I was able to successfully cross-foster three *T. g. castanotis* and two *T. g. guttata* at the egg stage. I successfully cross-fostered seven *T. g. castanotis* and six *T. g. guttata* at the fledgling stage. Time of fostering impacted significantly only two features of learned song. When all of the cross-fostered birds were combined together and divided by type of fostering, the birds that were cross-fostered at the egg stage were more similar (had a lower KL distance) to the tutors than those cross-fostered at the fledgling stage for 12 of the 13 features (Table 3.2, Figure 3.2). However, this was only significant for pitch (W = 5, p-value = 0.004435, Figure 3.2) and variance of entropy (W = 6, p-value = 0.007, Figure 3.2). I also combined the K-L measures across all of the features to see if the overall K-L distance differed significantly. This overall K-L measure was significant (W = 4496, p-value = 0.03, Figure 3.2).

Table 3.2. Means of K-L for the 13 features comparing cross-fostered zebra finches to tutors for egg cross-fostered vs. fledgling cross-fostered.

Means of K-L for the 13 features for birds cross-fostered at the egg stage compared to those cross-fostered at P30 compared to their tutors (Mann-Whitney Wilcoxon test, p-values < 0.05 denoted with *). Higher K-L means more variability.

	Amp	Pitch	FM	AM^2	Ent.	Pitch Good.	Mean Freq	Var. Pitch	Var. FM	Var. Ent.	Var. Pitch Good.	Var. Mean Freq	Var. AM
XF_Egg	3.2	1.87	2.29	2.14	3.34	2.51	3.56	0.41	1.39	1.02	0.85	0.87	2.04
XF_P30	4.33	4.23	2.99	2.58	4.28	3.98	2.24	0.6	1.83	2.64	1.36	1.73	2.56
p-value	*0.70*	*0.004**	*0.78*	*0.92*	*0.5*	*0.12*	*0.21*	*0.7*	*0.78*	*0.007**	*0.12*	*0.5*	*0.7*

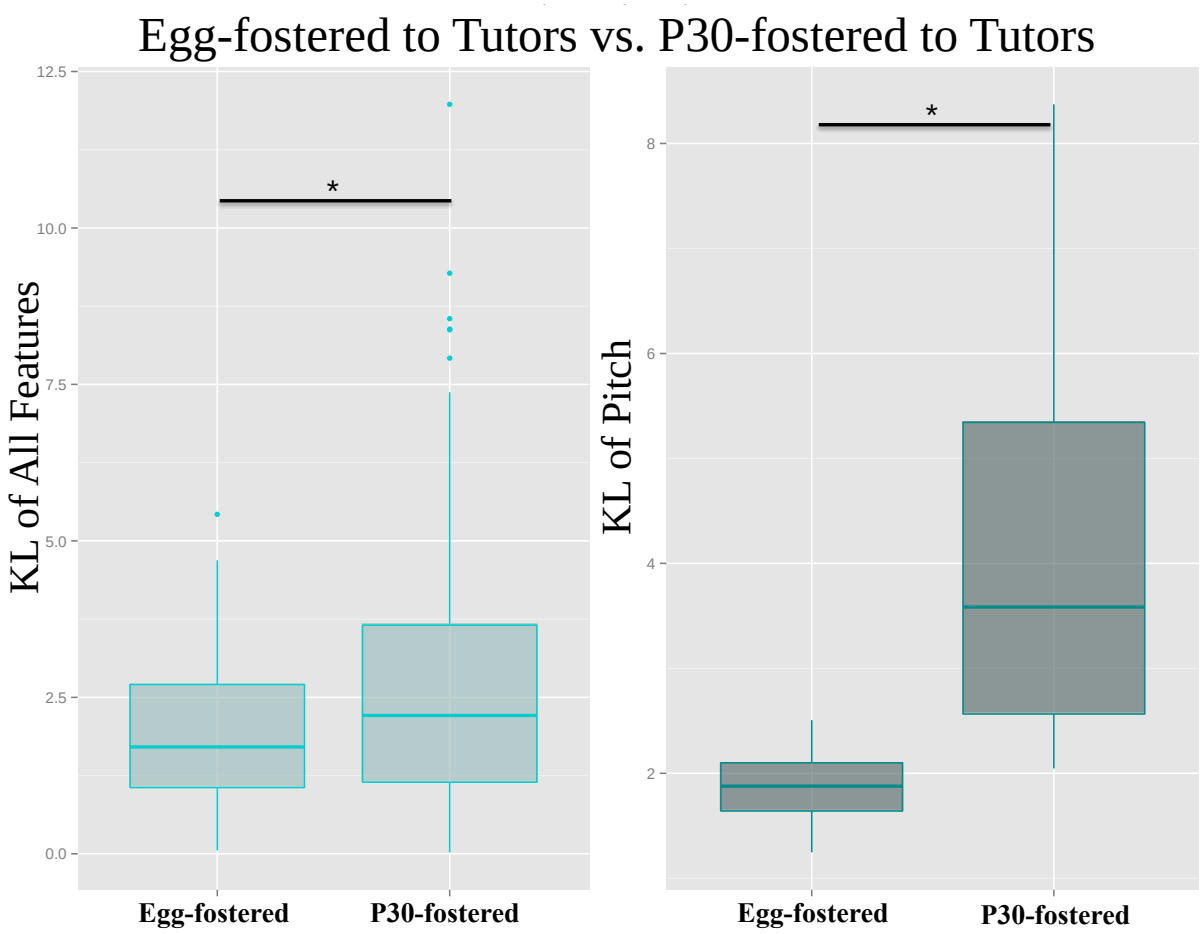

Figure 3.2. Variance of cross-fostered zebra finches to tutors for egg cross-fostered vs. fledgling cross-fostered.

The K-L distance for the zebra finches cross-fostered from the egg stage vs. those cross-fostered from the fledgling stage compared to their tutors for a) all features combined, and b) pitch. Black bars with an * denote that there is a significant difference between the two populations it connects.

Subspecies Differences in Song Tutor Copying

I quantified similarity between *T. g. castanotis*, *T. g. guttata* and their respective tutors using KL distance. When combining egg-fostered and P30-fostered birds, only one feature, variance of pitch, showed a statistically significant difference between the subspecies (W = 16, p-value = 0.03 Table 3.3) with *T. g. guttata* more closely matching the tutors. For all birds combined seven other features (amplitude, entropy, mean frequency, variance of pitch, variance of FM, variance of entropy, and variance of pitch goodness) trended in the same direction, and five features (FM, AM^2, pitch goodness, variance of mean frequency, and variance of AM) showed the opposite trend, but these differences were not statistically significant. For only the egg-fostered birds comparisons, *T. g. guttata* birds were more similar to the tutors for all features, but again, none were significant.

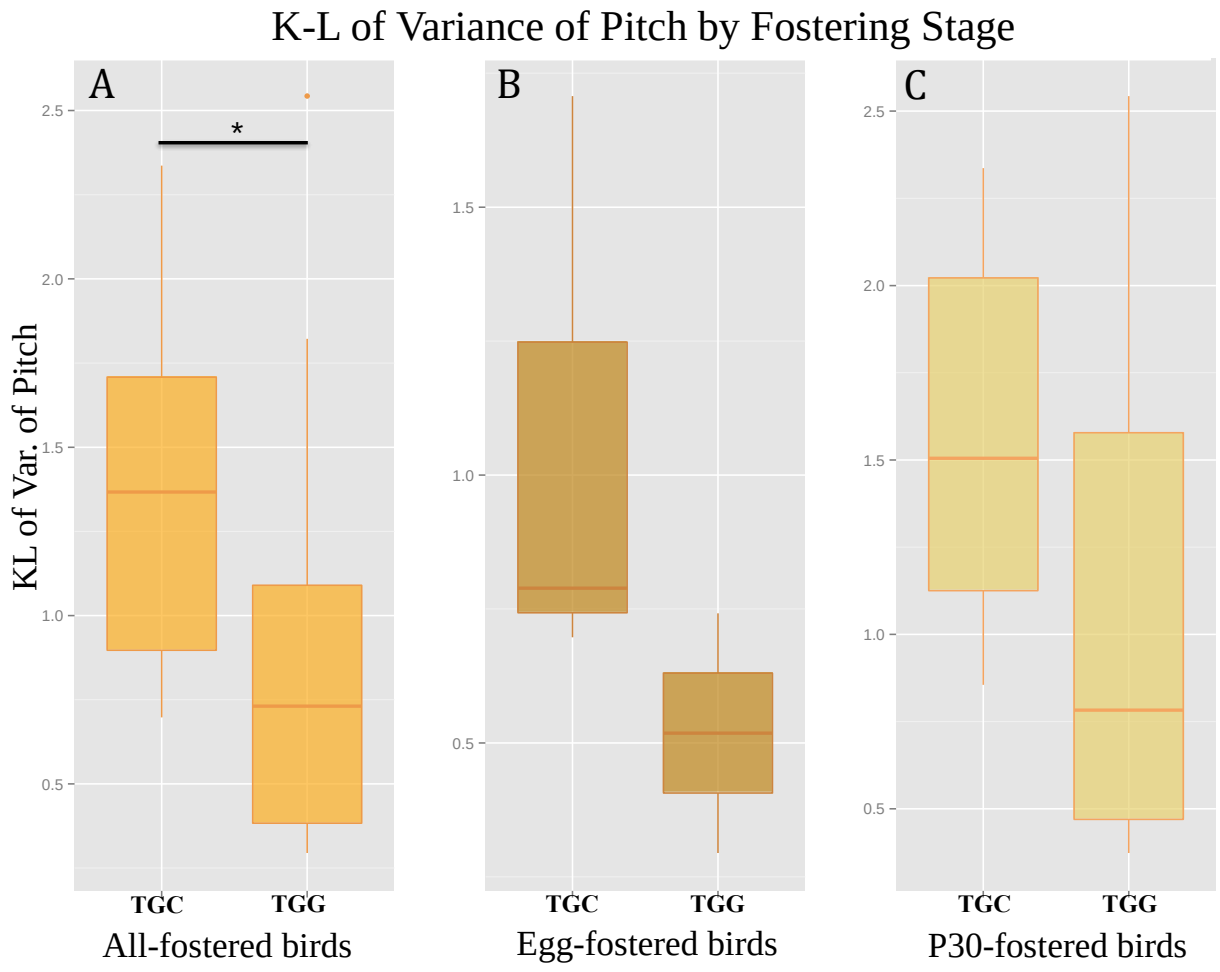

Figure 3.3. Variance of cross-fostered zebra finches to tutors for egg cross-fostered vs. fledgling cross-fostered by subspecies.

K-L distance for variance of pitch goodness of cross-fostered *T. g. castanotis* and *T. g. guttata* compared to their tutors for a) all the birds combined, b) only the egg-fostered birds, and c) only the fledgling fostered birds. Black bars with an * denote that there is a significant difference between the two populations it connects.

Table 3.3. Means of K-L for the 13 features comparing cross-fostered zebra finches to tutors for egg cross-fostered vs. fledgling cross-fostered by subspecies.

Means of K-L for the 13 features for the different fostering stages for cross-fostered *T. g. castanotis* vs. cross-fostered *T. g. guttata* compared to their tutors (Mann-Whitney Wilcoxon test, p-values < 0.05 denoted with *). Higher K-L means greater variability.

	Amp	Pitch	FM	AM^2	Ent.	Pitch Good.	Mean Freq	Var. Pitch	Var. FM	Var. Ent.	Var. Pitch Good.	Var. Mean Freq	Var. AM
XF_TGC to Tutor (All)	4.45	3.84	2.61	2.14	4.16	3.5	2.85	0.72	1.97	2.35	1.42	1.46	2.14
XF_TGG to Tutor (All)	3.49	3.25	3.02	2.86	3.83	3.65	2.29	0.33	1.38	1.99	0.97	1.53	2.77
p-value	*0.76*	*0.24*	*0.83*	*0.27*	*0.76*	*0.9*	*0.63*	*0.03**	*0.27*	*0.9*	*0.12*	*0.15*	*0.27*
XF_TGC to Tutor (Egg)	3.32	2.16	3.02	2.3	3.48	3.06	3.72	0.57	1.67	1.2	1.06	0.97	2.23
XF_TGG to Tutor (Egg)	3.05	1.44	1.19	1.91	3.14	1.68	3.32	0.17	0.98	0.76	0.52	0.73	1.76
p-value	*1*	*0.2*	*0.2*	*1*	*0.8*	*0.4*	*1*	*0.4*	*0.4*	*0.2*	*0.4*	*1*	*1*
XF_TGC to Tutor (Fledgling)	4.93	4.56	2.44	2.07	4.46	3.69	2.48	0.78	2.1	2.85	1.57	1.67	2.1
XF_TGG to Tutor (Fledgling)	3.64	3.85	3.63	3.18	4.06	4.31	1.95	0.39	1.51	2.41	1.12	1.8	3.11
p-value	*0.73*	*0.23*	*0.63*	*0.23*	*0.84*	*0.53*	*0.73*	*0.14*	*0.45*	*0.73*	*0.23*	*0.18*	*0.23*

In order to interpret song copying by zebra finches, I first described variability in tutors. In chapter 1 and 2, I showed that *T. g. castanotis* showed more individual variability in song than *T. g. guttata*. The Bengalese tutors used here were intermediate in variability, with an intermediate mean KL. The K-L for the within-Bengalese comparison was in between the K-L distance for the within-non-cross-fostered *T. g. castanotis* and the within-non-cross-fostered *T. g. guttata* birds. Bengalese finches are significantly less variable for all features when compared to the *T. g. castanotis* and are significantly more variable than the *T. g. guttata* birds for all features except variance of pitch and variance of pitch goodness (Table 3.4).

Table 3.4. Means of K-L for the 13 features comparing the two zebra finch subspecies to Bengalese finches.

Means of K-L for the 13 features for five Bengalese tutors and ten *T. g. castanotis* (TGC) males and ten *T. g. guttata* (TGG) males (chapter 1 and 2) (Mann-Whitney Wilcoxon test, p-values < 0.05 denoted with *). Higher K-L means greater variability.

	Amp	Pitch	FM	AM^2	Ent.	Pitch Good.	Mean Freq	Var. Pitch	Var. FM	Var. Ent.	Var. Pitch Good.	Var. Mean Freq	Var. AM
Bengalese	3.39	2.2	1.77	2.31	1.95	2.14	2.67	0.7	1.6	1.65	1.24	1.4	2.27
ECU_Dom_TGC	6.00	2.95	3.11	2.89	3.37	3.24	3.63	1.01	2.26	2.54	2.62	2.60	2.86
p-value	*3.53E-03**	*1.04E-02**	*1.05E-04**	*1.91E-02**	*1.35E-04**	*3.61E-03**	*2.86E-02**	*3.80E-03**	*2.35E-03**	*5.09E-04**	*2.62E-05**	*2.22E-04**	*1.48E-02**
Bengalese	3.39	2.2	1.77	2.31	1.94	2.14	2.67	0.7	1.6	1.65	1.24	1.4	2.27
ECU_TGG	1.95	1.18	0.93	0.96	0.99	0.94	1.01	0.27	0.55	0.76	0.76	0.73	0.95
p-value	*2.66E-04**	*7.62E-06**	*1.47E-03**	*1.20E-06**	*7.81E-04**	*1.95E-02**	*5.49E-09**	*1.46E-01*	*6.19E-08**	*1.53E-04**	*9.54E-01*	*3.28E-02**	*1.64E-06**

If zebra finch cross-fostering led to accurate copying of tutor songs, I might expect to see lower K-L between a bird and its tutor relative to comparisons with a non-tutor Bengalese finch. This is not what I observed. When I compared the cross-fostered birds to different Bengalese tutors other than their foster fathers, I did not see a significantly increased K-L for either subspecies (Table 3.5). For the cross-fostered *T. g. castanotis* individuals, there was a trend toward increased K-L when compared to Bengalese males that were not their tutors for 11 of the 13 features (all except FM and Pitch goodness), but similarity to the tutor was not significantly better than that to randomly selected birds. For the cross-fostered *T. g. guttata* individuals, there was no difference in song similarity with tutors than to randomly selected birds (Table 3.5). Given the moderate variability between Bengalese males (Table 3.4) instead of high variability seen between normally reared *T. g. castanotis* (chapter 1) for example, the fact that cross-fostered zebra finches are equally similar to non-tutor Bengalese finches may not necessarily mean that the birds were not learning Bengalese aspects of song.

Table 3.5. Means of K-L for the 13 features comparing cross-fostered zebra finches to tutors and to non-tutors by subspecies.

Means of K-L for the 13 features for the cross-fostered birds compared to their tutors vs. cross-fostered birds to randomly assigned non-tutors. (Mann-Whitney Wilcoxon test, p-values < 0.05 denoted with *). Higher K-L means greater variability.

	Amp	Pitch	FM	AM^2	Ent.	Pitch Good.	Mean Freq	Var. Pitch	Var. FM	Var. Ent.	Var. Pitch Good.	Var. Mean Freq	Var. AM
TGC to Tutor	4.45	3.84	2.61	2.14	4.16	3.5	2.85	0.72	1.97	2.35	1.42	1.46	2.14
TGC to Non-tutor	4.55	5.1	2.59	2.86	4.75	3.47	3.32	0.77	2.64	3.36	1.59	1.55	2.82
p-value	*0.8*	*0.11*	*0.91*	*0.17*	*0.58*	*0.97*	*0.63*	*0.44*	*0.25*	*0.08*	*0.53*	*0.48*	*0.25*
TGG to Tutor	3.49	3.25	3.02	2.86	3.83	3.65	2.3	0.33	1.38	1.99	0.97	1.53	2.77
TGG to Non-tutor	3.2	2.3	3.37	2.62	3.39	3.72	2.92	0.34	1.31	1.65	0.93	1.07	2.46
p-value	*0.8*	*0.51*	*0.51*	*1*	*0.44*	*0.88*	*0.72*	*0.38*	*0.8*	*0.65*	*0.72*	*0.38*	*0.72*

Subspecies Differences in Song Variability After Cross-fostering

Even if there is no difference between subspecies in terms of how accurately individuals copy their tutor, I also wanted to determine whether previously described song variability differences were maintained between subspecies after cross-fostering. For a subset of the cross-fostered birds that were balanced for which tutors they came from and which way they were cross-fostered, I found that this variability in *T. g. castanotis* individuals was maintained despite their being tutored by a different species (Figure 3.4, panels A-E).

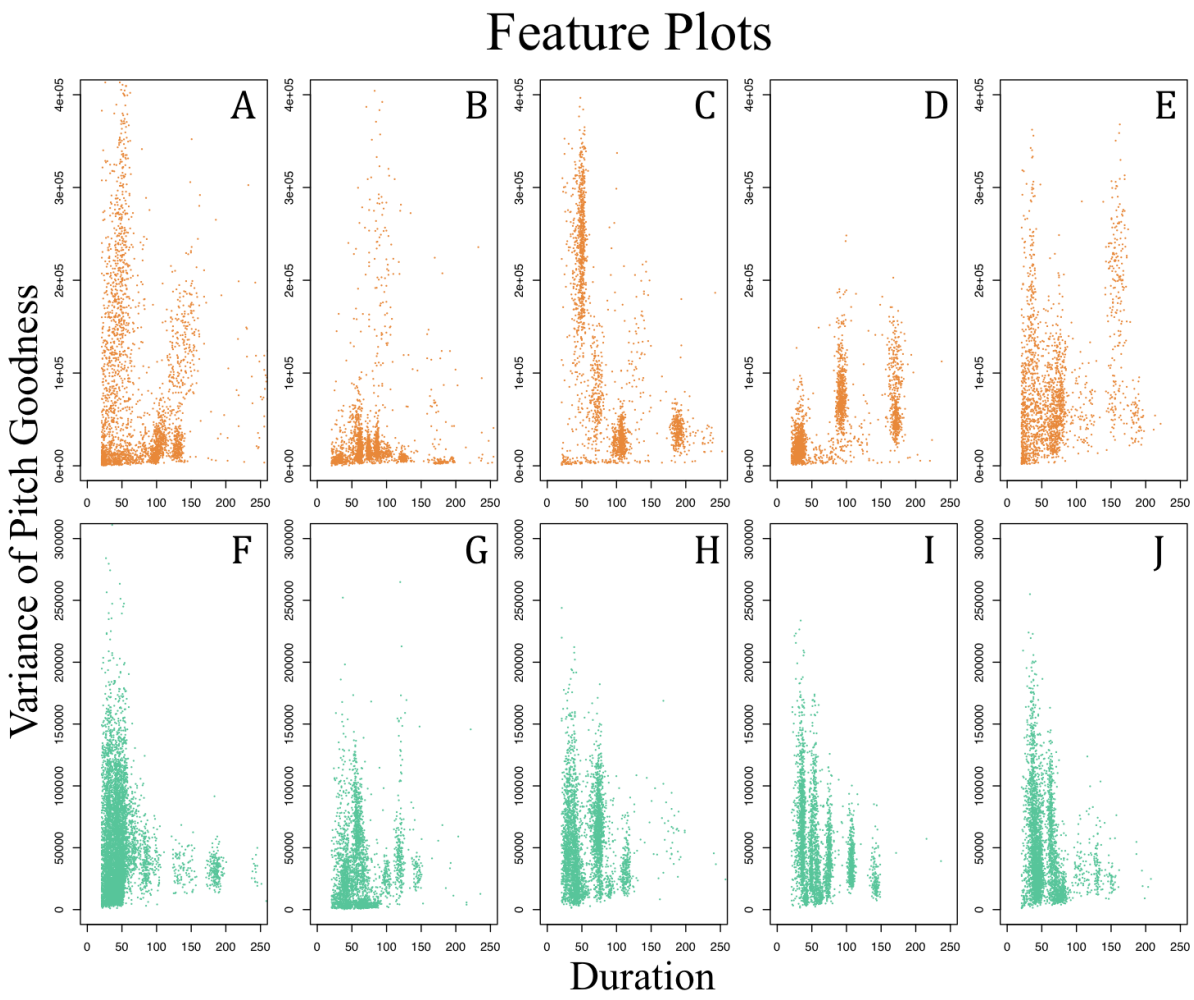

Figure 3.4. Features plots of duration vs. variance of pitch goodness for the 5 cross-fostered *T. g. castanotis* and 5 cross-fostered *T. g. guttata*.

Panels A-D represent five cross-fostered *T. g. castanotis*, one from each of the Bengalese tutors. Panels F-J represent five cross-fostered *T. g. guttata*, one from each of the same Bengalese tutors. Each feature plot shows duration vs. variance of pitch goodness. Within each parameter plot, a cluster of dots represents a syllable that is repeatedly sung. Therefore, it is possible to visually see the similarity in the overall pattern of clustered dots between each panel, which represents how similar the songs of each individual are to each other.

For example, for the *T. g. castanotis* in Figure 3.4, panel A and C have more note clusters in common to each other than A and D. Figure 5 panel B is a *T. g. castanotis* cross-fostered at the egg stage and the rest are cross-fostered at P30. Also following the pattern seen in chapter 2, the cross-fostered *T. g. guttata* birds showed a reduced variability between individuals (Figure 3.4, panels F-J). Panel C and panel E might be the most similar in terms of the clustering pattern, for example, but all *T. g. guttata* birds show very similar patterns of clustered notes for these features. Panel B is a *T. g. guttata* cross-fostered at the egg stage and the rest are cross-fostered at P30. For this balanced subset of cross-fostered birds, the *T. g. guttata* had statistically lower K-L for all features (Table 3.6, Figure 3.5), which means that the cross-fostered *T. g. guttata* showed less variability among individuals.

Table 3.6. Means of K-L for the 13 features comparing cross-fostered zebra finches to each other by subspecies.

Means of K-L for the 13 features for a within subspecies comparison of the cross-fostered birds (Mann-Whitney Wilcoxon test, p-values < 0.05 denoted with *). Higher K-L indicates more variability, which would mean that the cross-fostered *T. g. castanotis* are less similar to each other than *T. g. guttata* are.

	Amp	Pitch	FM	AM^2	Ent.	Pitch Good.	Mean Freq	Var. Pitch	Var. FM	Var. Ent.	Var. Pitch Good.	Var. Mean Freq	Var. AM
XF_TGC to Each Other	2.73	3.37	2.51	2.45	2.64	3.22	3.00	0.79	2.46	3.08	2.48	1.49	2.43
XF_TGG to Each Other	1.77	1.26	1.65	1.45	0.87	1.45	1.24	0.30	0.83	0.95	0.62	1.03	1.43
p-value	*0.005 *	*4.35e-10**	*0.001 *	*0.0008 *	*9.72e-10**	*1.34e-07**	*2.32e-08**	*3.36e-05**	*1.06e-07**	*1.02e-10**	*2.90e-11**	*0.007 *	*0.002 *

Figure 3.5. Variance of cross-fostered zebra finches to each other by subspecies for three features.

K-L distance for a) FM, b) Entropy, and c) variance of pitch goodness of five cross-fostered *T. g. castanotis* and five cross-fostered *T. g. guttata* compared to each other within each subspecies. Black bars with an * denote that there is a significant difference between the two populations it connects.

Discussion

Captive colonies of two zebra finch subspecies differ in overall song variability, with greater variability seen in *T. g. castanotis*. One possible cause of this difference is that *T. g. guttata* zebra finches more accurately copy their tutors. Song learning generates population variability in song structure (Zann 1993) and populations that are better at copying songs should have less variability than populations that are poor song copiers. To test for population differences in song copying, I cross-fostered the two subspecies of zebra finches to Bengalese finches and quantified song divergence between tutors and tutees and overall levels of variability. Interestingly, I found little evidence of population differences in song copying. Although the average K-L distance for eight of the thirteen features show a lower K-L (denoting more similarity) between cross-fostered *T. g. guttata* and tutors than the *T. g. castanotis* and their tutors, this difference is only statistically significant for one feature, variance of pitch (Table 3.3). Despite this, at the population level we still find higher variability in *T. g. castanotis* zebra finch songs suggesting that aspects of song production may contribute to the observed diversity of *T. g. castanotis* zebra finches' songs as is the case in naturally raised individuals (chapter 1). While I did not find an unambiguous difference in song copying, the maintenance of difference in variability under controlled tutoring is suggestive of a genetic difference in song production behavior or strong epigenetic/maternal effect.

Overall, the cross-fostered birds showed relatively low song similarity to their tutors. In fact, for the *T. g. guttata* especially, the cross-fostered birds were more similar to conspecifics for seven features than they were to their tutors. Additionally, similarity to the tutor was not significantly better than to randomly selected Bengalese for either subspecies. I found that Bengalese finches have a moderate degree of variability in song among individuals (more than *T.*

g. guttata and less than the highly variable *T. g. castanotis*). Perhaps moderate amount of variability explains why we could not see a difference between similarity to actual tutor and randomly selected Bengalese.

Although it appears that some amount of learning was taking place, perhaps since the degree of learning was rather small, it obscures my ability to quantify degree of similarity very accurately. The overall low similarity could reflect some of the same constraints on song production observed by Clayton (1990a-c). For all of the features combined and also for pitch specifically, I found that the birds that were cross-fostered at the egg stage (both subspecies combined) were more similar to their tutors; however, all other features independently showed no difference. If exposure to high diversity song during this time shapes subsequent song production then that might explain the difference observed. Both in the field and in lab experiments with song sparrows, it is has been shown that males learn from multiple tutors (Nordby et al. 2000; Beecher et al. 2007). Song tutoring throughout life resulted in significantly higher song similarity between tutor and tutee at least when looking across all features combined and for pitch specifically. It is possible that tutoring later in life, results in relatively poor song copying, and may have negatively impacted my ability to detect population differences in learning. Because of this, one caveat to the result of little similarity to tutors is that early developmental experience differed between treatments. Firstly, my sample size for egg-fostered birds is very small. Secondly, for the P30 individuals, each bird was raised with its own subspecies for 30 days, which includes ~10 days of the defined sensitive period for song learning. Though captive zebra finch generally learn from day 35-70 (Eales 1985; Clayton 1987), auditory learning in zebra finches begins around day 20, so even though they do not begin to practice their sub-song until day 35 (Wu et al. 2008), they will already have been learning their

conspecific songs for around 12 days before I cross-fostered them under the fledgling-fostered method. In particular, I missed a key part of the sensitive phase. In fact Clayton (1987) found that when the foster father after the sensitive phase is different from the foster father during that phase, learning of the tutor's song is not very successful. Only half of the birds raised with a zebra finch conspecific for the first 35 days and then a Bengalese finch tutor for days 35-70 learned some Bengalese song elements (Clayton 1987). My results appear to further support that conclusion.

I have also considered the possibility that certain aspects of song are affected differently by when the birds are exposed to a tutor. For example, since pitch was significantly more similar to tutors when cross-fostered from the egg stage it appears that it is not a physical constraint on matching the tutor's pitch; it is possible that certain features, such as pitch, are learned earlier than other features. Additionally, for the egg-fostered birds only comparison, *T. g. guttata* birds were more similar to the tutors for all features, though none were significant. This could be due to the small sample size of birds fostered in this manner however. In the future it would be useful to attempt to obtain more egg-fostered zebra finches, despite the challenges I encountered with this approach.

As mentioned, another limitation in this study was sample size. Although many statistics trend towards *T. g. guttata* zebra finches showing higher copying, poor learning of tutor song at P30, and high variance yielded limited power to detect significant differences. While at present I cannot reject the null hypothesis that zebra finches in this study copy equally, it does appear that the greater variability among *T. g. castanotis* individuals remains whether they are cross-fostered to a different species or not. An alternative approach to examine these questions would be to study song copying in subspecies and their F1 hybrids. Quantitative genetic approaches have

now been used successfully to map genetic underpinnings of behavioral traits in numerous species (Ding et al. 2016; Fergus and Shaw 2011; Greenwood et al. 2013; Shirang et al. 2013; Weber et al. 2013), and such approaches may hold promise for zebra finches as well.

References

Airey DC, Castillo-Juarez H, Casella G, Pollak EJ, DeVoogd TJ (2000) Variation in the volume of zebra finch song control nuclei is heritable: Developmental and evolutionary implications. *Proceedings of the Royal Society B: Biological Sciences,* 267:2099-2104.

Anderson, M, Iwasa Y (1996) Sexual selection. *Trends Ecol Evol,* 11:53–58.

Balakrishnan CN, Edwards SV (2009) Nucleotide variation, linkage disequilibrium and founder-facilitated speciation in wild populations of the zebra finch (*Taeniopygia guttata*). *Genetics,* 181: 645-660.

Beecher MD, Burt JM, O'Loghlen AL, Templeton CN, Campbell SE (2007) Bird song learning in an eavesdropping context. *Anim Behav,* 73:929–935.

Boogert NJ, Giraldeau LA, Lefebvre L (2008) Song complexity correlates with learning ability in zebra finch males. *Anim Behav* 76:1735–41.

Buchanan KL, Catchpole CK (2000) Song as an indicator of male parental effort in the sedge warbler. *Proc Biol Sci,* 267:321–326.

Buchanan KL, Leitner S, Spencer KA, Goldsmith AR, Catchpole CK (2004) Developmental stress selectively affects the song control nucleus HVC in the zebra finch. *Proc R Soc Lond B,* 271:2381–2386

Buchanan KL, Leitner S, Spencer KA, Goldsmith AR, Catchpole CK (2004) Developmental stress selectively affects the song control nucleus HVC in the zebra finch. *Proc Biol Sci,* 271:2381–2386.

Burkett ZD, Day NF, Peñagarikano O, Geschwind DH, White SA (2015) VoICE: A semi-automated pipeline for standardizing vocal analysis across models. *Scientific Reports* 5.

Campbell DLM, Hauber ME (2009) The disassociation of visual and acoustic conspecific cues decreases discrimination by female zebra finches (Taeniopygia guttata). *Journal Of Comparative Psychology,* 123(3):310-315.

Campbell DLM, Hauber ME (2011) Cross-fostering diminishes song discrimination in zebra finches (Taeniopygia guttata). *Animal Cognition,* 12:481-90.

Campbell DLM, Shaw RC, Hauber ME (2009) The Strength of Species Recognition in Captive Female Zebra Finches (Taeniopygia guttata): A Comparison Across Estrildid Heterospecifics. *Ethology,* 115:23–32.

Catchpole CK (1986) Song repertoires and reproductive success in the great reed warbler Acrocephalus arundinaceus. *Behav Ecol Sociobiol,* 19:439–445.

Catchpole CK (1980) Sexual selection and the evolution of complex songs among European warblers of the genus Acrocephalus. *Behaviour,* 74:149–166.

Catchpole CK (1996) Female aquatic warblers (Acrocephalus paludicola) are attracted by playback of longer and more complicated songs. *Behavior,* 133:15–16.

Clayton NS (1987). Song learning in cross-fostered zebra finches: a re-examination of the sensitive phase. *Behaviour,* 102:6~81.

Clayton NS (1989) The effects of cross-fostering on selective song learning in estrildid finches. *Behaviour,* 190:163-74.

Clayton NS (1990a) Assortative Mating in Zebra Finch Subspecies, Taeniopygia guttata guttata and *T. g.castanotis*. *Philosophical Transactions: Biological Sciences,* 330(1258):351-370.

Clayton NS (1990b) Mate choice and pair formation in Timor and Australian Mainland zebra finches. *Anim Behav,* 39:474-480.

Clayton NS (1990c) Subspecies recognition and song learning in zebra finches. *Anim Behav,* 40:1009-1017.

Dave AS, Margoliash D (2000) Song Replay During Sleep and Computational Rules for Sensorimotor Vocal Learning. *Science,* 290(5492):812-816.

Davidson JD, Balakrishnan CN (2016) Gene regulatory divergence during speciation in a songbird. *G3: Genes, Genomes & Genetics,* 6:1357-1364.

Deacon TW (2010) Colloquium paper: a role for relaxed selection in the evolution of the language capacity. *Proc Natl Acad Sci USA.,* 107:9000–9006.

Derégnaucourt S, Mitra PP, Fehér O, Pytte C, Tchernichovski O (2005) How sleep affects the developmental learning of bird song. *Nature,* 433:710–6.

Ding Y, Berrocal A, Morita T, Longden KD, Stern DL (2016) Natural courtship song variation caused by an intronic retroelement in an ion channel gene. *Nature,* 536(7616):329-332.

Doupe AJ, Kuhl PK (1999) Birdsong and human speech: common themes and mechanisms. *Annual Review of Neuroscience,* 22:567-631.

Eales LA (1985) Song learning in zebra finches: some effects of song model availability on what is learnt and when. *Animal Behaviour, 33*(4):1293-1300.

Eales LA (1987a) Do zebra finch males that have been raised by another species still tend to select a conspecific song tutor? *Anim Behav,* 35**:**1347–1355.

Eales LA (1987b) Song learning in female-raised zebra finches: another look at the sensitive phase. *Animal Behaviour,* 35:1356–1365.

Fehér O, Wang H, Saar S, Mitra PP, Tchernichovski O (2009) De novo establishment of wild-type song culture in the zebra finch. *Nature,* 459:564-568.

Fergus DJ, Shaw KL (2011) Genetically regulated temporal variation of novel courtship elements in the Hawaiian cricket genus Laupala. *Behavior genetics,* 41(4):607-614.

Forstmeier W, Segelbacher G, Mueller JC, Kempenaers B (2007) Genetic variation and differentiation in captive and wild zebra finches (*Taeniopygia guttata*). *Molecular ecology,* 16:4039–50.

Forstmeier W, Burger C, Temnow K, Derégnaucourt S (2009) The genetic basis of zebra finch vocalizations. *Evolution,* 63:2114-2130.

Gentner TQ, Hulse SH (2000) Female European starling preference and choice for variation in conspecific male song. *Anim Behav,* 59:443–458.

Gil D, Gahr M (2002) The honesty of bird song: multiple constraints for multiple traits. *Trends Ecol Evol,* 17:133–141.

Grafen A (1990) Biological signals as handicaps. *J Theor Biol,* 144:517–546.

Griffiths R, Double MC, Orr K, Dawson RJG (1998) A DNA test to sex most birds. *Molecular,* 7:1071–1075.

Greenwood AK, Wark AR, Yoshida K, Peichel CL (2013) Genetic and neural modularity underlie the evolution of schooling behavior in threespine sticklebacks. *Current Biology,* 23(19): 1884-1888.

Hasselquist D, Bensch S, von Schantz T (1996) Correlation between male song repertoire, extra-pair paternity and offspring survival in the great reed warbler. *Nature,* 381:229–232.

Hawkins CJ (1918) Sexual Selection and Bird Song. *The Auk,* 35(4):421-437.

Hiebert SM, Stoddard PK, Arcese P (1989) Repertoire size, territory acquisition and reproductive success in the song sparrow. *Anim Behav,* 37:266–273.

Honda E, Okanoya K. Acoustical and Syntactical Comparisons between Songs of the White-backed Munia (*Lonchura striata*) and Its Domesticated Strain, the Bengalese Finch (*Lonchura striata var. domestica*). *Zoological Science,* 16:319–326.

Jensen T, Pernasetti FMM, Durrant B (2003) Conditions for rapid sex determination in 47 avian species by PCR of genomic DNA from blood, shell-membrane blood vessels and feathers. *Zoo boil,* 22:561-571.

Jin H, Clayton DF (1997) Localized Changes in Immediate-Early Gene Regulation during Sensory and Motor Learning in Zebra Finches. *Neuron,* 19:1049–1059.

Johnstone RA, Grafen A (1993) Dishonesty and the handicap principle. *Anim Behav,* 46:759–764.

Kagawa H, Suzuki K, Takahasi M, Okanoya K (2014) Domestication changes innate constraints for birdsong learning. *Behavioural Processes,* 106:91–97.

Khan NW, St John J, Quinn TW (1998) Chromosome-specific intron size differences in the avian CHD gene provides an efficient method for sex identification in birds. *The Auk,* 155:1074-1078.

Kolts JR, McRae SB (2017) Seasonal home range dynamics and sex differences in habitat use in a threatened, coastal marsh bird. *Ecology and Evolution*, 7(4):1101-1111.

Kroodsma D (1976) Reproductive development in a female songbird: differential stimulation by quality of male song. *Science,* 192:574–575.

Lampe HM, Saetre GP (1995) Female pied flycatchers prefer males with larger song repertoires. *Proc Biol Sci,* 262:163–167.

Leitão A, ten Cate C, Riebel K (2006) Within-song complexity in a songbird is meaningful to both male and female receivers. *Anim Behav,* 71:1289–1296.

London SE, Clayton DF (2008) Functional identification of sensory mechanisms required for developmental song learning. *Nature Neuroscience,* 11:579-586.

Marler P (1970) Birdsong and speech development: could there be parallels? *American Scientist,* 58:669-73.

Marler P, Peters S (1977) Selective vocal learning in a sparrow. *Science,* 198,519–521.

Marler P, Tamura M (1962) Song "Dialects" in Three Populations of White-Crowned Sparrows. *The Condor*, 64(5):368-377.

Mayr E (1944) Timor and the colonization of Australia by birds. *Emu,* 44:113-30.

Mello C (2014) The Zebra Finch, *Taeniopygia guttata*: An Avian Model for Investigating the Neurobiological Basis of Vocal Learning. *Cold Spring Harbor Protocols,* 12:1237-1242.

Mooney R (2009) Neural mechanisms for learned birdsong. *Learning and Memory,* 16:655–69.

Mori C, Wada K (2015). Audition-Independent Vocal Crystallization Associated with Intrinsic Developmental Gene Expression Dynamics. *Journal of Neuroscience,* 35(3):878-889.

Nelson DA (2000) Song overproduction, selective attrition and song dialects in the white-crowned sparrow. *Animal Behaviour* 60(6), 887-898.

Newhouse DJ, Balakrishnan CN (2015) High MHC diversity despite bottlenecks in wild and domesticated zebra finches. *BMC Evolutionary Biology,* 15:256.

Nordby JC, Campbell SE, Burt JM, Beecher MD (2000) Social influences during song development in the song sparrow: a laboratory experiment simulating field conditions. *Anim Behav,* 59:1187–1197.

Nottebohm F, Kasparian S, Pandazis C (1981) Brain space for a learned task. *Brain Res,* 213:99–109.

Nowicki S, Searcy WA, Peters S (2002) Brain development, song learning and mate choice in birds: a review and experimental test of the "nutritional stress hypothesis". *J Comp Physiol A,* 188:1003–1014.

Ohgushi E, Mori C, Wada K (2015) Diurnal oscillation of vocal development associated with clustered singing by juvenile songbirds. *Journal of Experimental Biology,* 218:2260-2268.

Okanoya K (2002) Sexual display as a syntactical vehicle: the evolution of syntax in birdsong and human language through sexual selection. *The transition to language,* 2:46-63.

Okanoya K (2004a) Song syntax in Bengalese finches: proximate and ultimate analyses. *Adv Stud Behav,* 34:297–346.

Okanoya K (2004b) The Bengalese finch: a window on the behavioral neurobiology of birdsong syntax. *Ann NY Acad Sci,* 1016:724–735.

Okanoya K (2012) Behavioural Factors Governing Song Complexity in Bengalese Finches. *International Journal of Comparative Psychology,* 25:44-59.

Olveczky BP, Andalman AS, Fee MS (2005) Vocal experimentation in the juvenile songbird requires a basal ganglia circuit. *PLoS Biology,* 3:e153.

Peters S, Searcy WA, Nowiki S (2014) Developmental Stress, Song-Learning, and Cognition. *Integrative and Comparative Biology*, 54(4):555–567.

Podos J, Huber SK, Taft B (2004a) Bird song: The interface of evolution and mechanism. *Ann NY Acad Sci,* 35:55–87.

Price T (2008) *Speciation in birds*. Roberts and Company Publishers.

R Core Team (2014) R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Reid JM, Arcese P, Cassidy A, Hiebert SM, Smith JNM, Stoddard PK, Marr AB, Keller LF (2004) Song repertoire size predicts initial mating success in male song sparrows, Melospiza melodia. *Anim Behav,* 68:1055–63.

Rogers CH (1979) Zebra finches. K & R Books.

Searcy WA (1992) Song Repertoire and Mate Choice in Birds. *Integr Comp Biol,* 32(1):71-80.

Shirangi TR, Stern DL, Truman JW (2013) Motor control of Drosophila courtship song. *Cell reports,* 5(3):678-686.

Simonyan K, Horwitz B, Jarvis ED (2012) Dopamine regulation of human speech and bird song: A critical review. *Brain & Language,* 122:142–150.

Slater PJB, Eales LA, Clayton NS (1988) Song learning in zebra finches (Taeniopygia guttata): progress and prospects. *Advances in the Study of Behavior,* 18:1-34.

Sound Analysis Pro http://ofer.sci.ccny.cuny.edu/html/sound_analysis.html

Soma MF (2011) Social Factors in Song Learning: A Review of Estrildid Finch Research. *Ornithological Science,* 10(2):89-100.

Suzuki K, Ikebuchi M, Bischof H, Okanoyaa K (2013) The impact of domestication on fearfulness: A comparison of tonic immobility reactions in wild and domesticated finches. *Behavioural Processes,* 100:58– 63.

Suzuki K, Ikebuchi M, Bischof H, Okanoyaa K (2014) Behavioral and neural trade-offs between song complexity and stress reaction in a wild and a domesticated finch strain. *Neuroscience and Biobehavioral Reviews,* 46:547–556.

Suzuki K, Yamada H, Kobayashi T, Okanoya K (2012) Decreased fecal corticosterone levels due to domestication: a comparison between the white-backed munia (*Lonchura striata*) and its domesticated strain, the Bengalese finch *(Lonchura striata var. domestica*) with a suggestion for complex song evolution. *J Exp Zool,* 317A:561–570.

Takahasi M, Kagawa H, Ikebuchi M, Okaoya K (2006) Case studies of song and call learning by a hybrid Bengalese–Zebra Finch and Bengalese-fostered Zebra Finches: assessing innate factors in vocal learning. *Ornithol Sci,* 5:85–93.

Tchernichovski O, Nottebohm F, Ho CE, Pesaran B, Mitra PP (2000) A procedure for an automated measurement of song similarity. *Animal Behaviour,* 59:1167–1176.

Tchernichovski O, Mitra PP, Lints T, and Nottebohm F (2001) Dynamics of the vocal imitation process: how a zebra finch learns its song. *Science,* 291:2564–9.

ten Cate C (1987) Sexual preferences in zebra finch males raised by two species: II. The internal representation resulting from double imprinting. *Anim Behav*, 35:321–330.

ten Cate C (2014) On the phonetic and syntactic processing abilities of birds: From songs to speech and artificial grammars. *Current Opinion in Neurobiology*, 28:157–164.

Thompson JA, Basista MJ, Wu W, Bertram R, Johnson F (2011) Dual pre-motor contribution to songbird syllable variation. *The Journal of Neuroscience*, 31:322–330.

Vallentin D, Kosche G, Lipkind D, Long MA (2016) Inhibition protects acquired song segments during vocal learning in zebra finches. *Science,* 351(6270):267-271.

Villain AS, Boucaud ICA, Bouchut C, Vignal C (2015) Parental influence on begging call structure in zebra finches (*Taeniopygia guttata*): evidence of early vocal plasticity. *R Soc Open Sci,* 2:150497.

Weber JN, Peterson BK, Hoekstra HE (2013) Discrete genetic modules are responsible for complex burrow evolution in Peromyscus mice. *Nature,* 493(7432):402-405.

Woodgate JL, Buchanan KL, Bennett ATD, Catchpole CK, Brighton R, Leitner S (2014) Environmental and genetic control of brain and song structure in the zebra finch. *Evolution,* 68:230-240.

Woolley SC, Doupe AJ (2008) Social context-induced song variation affects female behavior and gene expression. *PLoSBiol,* 6(3):e62.

Wu W, J. Thompson JA, Bertram R, Johnson F (2008) A statistical method for quantifying songbird phonology and syntax. *Journal of Neuroscience Methods,* 174:147–154.

Zahavi A (1975) Mate selection — a selection for a handicap. *J Theor Biol* 53, 205–214.

Zann RA (1983) Structural Variation in the Zebra Finch Distance Call. *Zeitschrift für Tierpsychologie,* 66(4):328–345.

Zann RA (1993) Variation in song structure within and among populations of Australian zebra finches. *The Auk*, 110(4):716-726.

Zann RA (1996) *The zebra finch: a synthesis of field and laboratory studies*. Oxford: Oxford University Press.

"Zebra Finch - *Taeniopygia guttata castanotis*." http://www.efinch.com/species/zebra.html.

CHAPTER 4 BEHAVIORAL AND GENETIC DIVERGENCE AMONG WILD AND DOMESTICATED POPULATIONS OF THE ZEBRA FINCH (*TAENIOPYGIA GUTTATA*): A SYNTHESIS

Summary

Here I present the first genomic analysis of the zebra finch in the context of speciation and domestication. I have taken a first step to characterize the genomic and behavioral variation in this system, which is the initial step to being better able to study song learning in a genetics context. In both analyses I see an average genome wide F_{ST} of around 0.02, but in both domestication and speciation scenarios there are regions in the genome that are highly diverged in the populations. There are far more of these outlier regions in the subspecies comparison than the domestication comparison. I have found a reduction of diversity in both the domesticated population as well as the island populations despite widely varying timescales (~150 years and 1-2MYA) (Balakrishnan and Edwards 2009).

I find that the historical bottleneck in the island subspecies, *T. g. guttata*, also likely has affected aspects of behavior, specifically song production. The island subspecies shows less variation in the songs produced among individuals than the mainland birds, *T. g. castanotis*. The domestication process has had the opposite result: I observed an increased variability in song structure in domesticated birds. One explanation for this pattern is that domestication has freed captive populations from the constraints faced in the wild due to mate choice and the need for accurate species recognition. Through cross-fostering I took the first step to differentiate between genetic or cultural controls of this difference in variation. I did not find a difference in the song-copying behavior of the two subspecies when they were cross-fostered to the Bengalese finch.

However, the high variability in song structure in *T. g. castanotis* remains following controlled tutoring, which might indicate a genetic role.

Constraints/future changes to study design

My study had a number of constraints the biggest being with regards to the song learning work. Since I was unable to obtain a large enough sample size of cross-fostered birds from the egg stage, I was limited in the amount of learning that I could detect. Zebra finches that were cross-fostered at P30 spent around ten days with conspecifics during the start of sensory period (Clayton 1987). I believe that it is important to continue trying to cross-foster from the egg stage to eliminate this complication. Additionally, though visually the tutors and tutees were blocked from each other, all the cross-fostering took place in one room that also contained the flight aviaries with adult *T. g. castanotis* and *T. g. guttata*. If this study is continued in the future I suggest creating more isolation between the cross-fostering set-ups and the rest of the zebra finches. Additionally, it will be useful to run a mixed-model on the song data to separate effect of population and the specific combination of birds that are being compared.

I think that it is also important to sample another few domesticated zebra finch populations to better support the conclusions. East Carolina University's population of *T. g. castanotis* were more variable among individuals than University of Chicago's were, so it would be useful to expand beyond solely including these two populations. It would also be interesting to include some populations of domesticated *T. g. castanotis* from Europe since the birds bred in captivity in each continent have been isolated from each other (Forstmeier et al. 2007). It would be useful to see if this increased variability in songs produced has resulted separately in the European populations as well because this would lend further support to the role domestication

plays in producing this greater variability. Also, if increased variability is a characteristic of the European populations as well, we could then see if the underlying genetic changes differed despite producing the same behavioral result.

As for the genomics work, one of the major goals going forward is to better tease apart the influences of selection and drift. I was already able to identify regions of selection based on linkage disequilibrium using the program OmegaPlus in the domestication analyses, but I would like to do the same in the speciation analysis. Additionally, it will be important to model neutral evolution in this scenario so we can identify what changes might have been impacted by selection. I also have the opportunity to estimate demographic histories using PSMC or MSMC (PSMC - http://github.com/lh3/psmc). For the study of speciation specifically, it would be interesting to do other types of demographic analyses to see if there is evidence for more a recent bottleneck that the colonization of the islands 1-2MYA (Balakrishnan and Edwards 2009). We could estimate how old the novel lineages for different alleles are and when they started to diverge within the population. I also would like to test to see if our areas of high F_{ST} fall in areas of low recombination using Singhal and colleagues' (2015) data on linkage-disequilibrium in zebra finches. Additionally, there is also an updated zebra finch genome coming out (Korlach et al. 2017) that will hopefully be better annotated at some of the outlier sites that I have identified. I am also interested if these outlier sites are within constrained regions. With the sites within the pigmentation genes in particular, I want to get more information about these SNPs to see if they are worth investigating in a functional study in the future.

Overall, I have taken the first step to characterize the intraspecific genetic and behavioral variation in wild and domesticated populations of zebra finches. Having this understanding of the

variation in the system will allow us to study gene-environment influences on behavior and specifically investigate potential genetic contributions to song copying ability.

References

Balakrishnan CN, Edwards SV (2009) Nucleotide variation, linkage disequilibrium and founder-facilitated speciation in wild populations of the zebra finch (*Taeniopygia guttata*). *Genetics,* 181: 645-660.

Clayton NS (1987). Song learning in cross-fostered zebra finches: a re-examination of the sensitive phase. Behaviour, 102:6~81.

Forstmeier W, Segelbacher G, Mueller JC, Kempenaers B (2007) Genetic variation and differentiation in captive and wild zebra finches (*Taeniopygia guttata*). *Molecular ecology* 16:4039–50.

Korlach J, Gedman G, Kingan S, Chin J, Howard J, Cantin L, Jarvis ED (2017) De Novo PacBio Long-Read and Phased Avian Genome Assemblies Correct and Add to Genes Important in Neuroscience Research. *bioRxiv*, http://biorxiv.org/content/early/2017/02/02/103911.abstract.

PSMC - http://github.com/lh3/psmc

Singhal S, Leffler EM, Sannareddy K, Turner I, Venn O, Hooper DM, Strand AI, Li Q, Raney B, Balakrishnan CN, Griffith SC, McVean G, Przeworski M (2015) Stable recombination hotspots in birds. *Science,* 350(6263):928-932.

APPENDIX A: ANIMIAL USE PROTOCOL (AUP)

Table A.1. Animal use protocol (AUP) for the work presented in this dissertation.

East Carolina University's AUP number designations with the chapters in which the AUP was utilized.

AUP #	Organization	Applicable Chapters	Page #
D278a	East Carolina University	Chapters 1-2	126
D285a	East Carolina University	Chapter 3	127

East Carolina University

Animal Care and Use Commitee
212 Ed Warren Life Sciences Building
East Carolina University
Greenville, NC 27834

252-744-2436 office
252-744-2355 fax

May 12, 2015

Chris Balakrishnan, Ph.D.
Department of Biology
Howell Science Complex
East Carolina University

Dear Dr. Balakrishnan:

Your Animal Use Protocol entitled, "Breeding and Holding Protocol for Finches" (AUP #D278a) was reviewed by this institution's Animal Care and Use Committee on May 11, 2015. The following action was taken by the Committee:

"Approved as submitted"

Please contact Dale Aycock at 744-2997 prior to hazard use

A copy is enclosed for your laboratory files. Please be reminded that all animal procedures must be conducted as described in the approved Animal Use Protocol. Modifications of these procedures cannot be performed without prior approval of the ACUC. The Animal Welfare Act and Public Health Service Guidelines require the ACUC to suspend activities not in accordance with approved procedures and report such activities to the responsible University Official (Vice Chancellor for Health Sciences or Vice Chancellor for Academic Affairs) and appropriate federal Agencies. **Please ensure that all personnel associated with this protocol have access to this approved copy of the AUP and are familiar with its contents.**

Sincerely yours,

Susan McRae, Ph.D.
Chair, Animal Care and Use Committee

SM/jd

enclosure

East Carolina University is a constituent institution of the University of North Carolina. An equal opportunity university.

East Carolina University

Animal Care and Use Commitee
212 Ed Warren Life Sciences Building
East Carolina University
Greenville, NC 27834

252-744-2436 office
252-744-2355 fax

September 15, 2015

Christopher Balakrishnan, Ph.D.
Department of Biology
Howell Science Complex
East Carolina University

Dear Dr. Balakrishnan:

Your Animal Use Protocol entitled, "Gene Expression and Social Behavior in Zebra Finches" (AUP #D285a) was reviewed by this institution's Animal Care and Use Committee on September 14, 2015. The following action was taken by the Committee:

"Approved as submitted"

Please contact Dale Aycock at 744-2997 prior to hazard use

A copy is enclosed for your laboratory files. Please be reminded that all animal procedures must be conducted as described in the approved Animal Use Protocol. Modifications of these procedures cannot be performed without prior approval of the ACUC. The Animal Welfare Act and Public Health Service Guidelines require the ACUC to suspend activities not in accordance with approved procedures and report such activities to the responsible University Official (Vice Chancellor for Health Sciences or Vice Chancellor for Academic Affairs) and appropriate federal Agencies. **Please ensure that all personnel associated with this protocol have access to this approved copy of the AUP and are familiar with its contents.**

Sincerely yours,

Susan McRae, Ph.D.
Chair, Animal Care and Use Committee

SM/jd

enclosure

East Carolina University is a constituent institution of the University of North Carolina. An equal opportunity university.