Analysis of female song provides insight into the evolution of dimorphism in a widely studied songbird

Matthew R. Wilkins[1,2], Karan J. Odom[3], Lauryn Benedict[4], Rebecca J. Safran[1]

Keywords: sexual selection, female song, complex signaling, multimodal signaling, *Hirundo rustica*, animal communication

1-University of Colorado at Boulder, Department of Ecology and Evolutionary Biology

2-Vanderbilt University, Center for Science Outreach

3-Cornell University, Cornell Lab of Ornithology and Department of Neurobiology and Behavior

4-University of Northern Colorado, School of Biological Sciences

ABSTRACT

Understanding the patterns and processes related to sexual dimorphism in diverse animal taxa is a foundational research topic in ecology and evolution. Within the realm of animal communication, studies have traditionally focused on male signals, assuming that female choice and male-male competition have promoted dimorphism via elaboration of male traits, but selection on females also has the potential to create sex differences. Here, we describe female song in barn swallows for the first time, report rates of female song production, and couple song data with plumage data to explore the relative degree to which dimorphism in signaling traits is consistent with contemporary selection on males versus females. During previous intensive study of male song over two years, we recorded songs for 15 females, with matched phenotypic and fitness data. We randomly selected 15 samples from our larger male dataset to test whether sexual dimorphism in song and plumage is more strongly associated with fledgling success for females or genetic paternity for males. Analyses included 35 potential sexual signals including 22 song parameters and 13 plumage traits. Outcomes indicate that: female songs are used in multiple contexts, restricted primarily to the beginning of the breeding season; song traits were more dimorphic than visual plumage traits; and trait correlations with reproductive success in females, rather than males, predicted sex-based differences in song and plumage. These results are consistent with phylogenetic studies showing that sex-based phenotypic differences are driven by changes in females, highlighting the potential role of female trait evolution in explaining patterns of sexual dimorphism. To achieve a better understanding of dimorphism, we require comprehensive studies that measure the same traits in males and females and their fitness consequences.

INTRODUCTION:

Sexual dimorphism and the processes driving it have long been a central research topic in ecology and evolution (Andersson, 1994; Alexander V. Badyaev & Hill, 2003; Burns, 1998; C. Darwin, 1859, p. 94; E. Darwin, 1801, p. 396; Hedrick & Temeles, 1989; Lande, 1980; Ng et al., 2019). However, owing to historical biases, studies of the drivers of dimorphism in sexual signaling traits have traditionally focused on male signals, and most approaches assume that sexual selection has promoted dimorphism via elaboration of male traits (Alexander V. Badyaev & Hill, 2003; Freed, 2000; Langmore, 1998; Riebel, 2016; Riebel et al., 2005, 2019; Rosvall, 2011). Yet, meta-analyses of the strength of sexual selection on male traits report moderate effect sizes (Jennions et al., 2012), with evidence that males are often under variable selection

pressures within and across breeding seasons (Chaine & Lyon, 2008; Kingsolver et al., 2012; Robinson et al., 2008; Steele et al., 2011), have trait values near optima (Evans, 1998; Rodríguez et al., 2006), or that the magnitude of trait dimorphism may not reflect the strength of current selection on males (Miller et al., 2016). Indeed, dimorphism in signals can be caused by a range of selection pressures on both males and females, resulting in exaggeration or reduction of a variety of sex-specific signals (Bell & Zamudio, 2012a; Price, 2015a; Wiens, 2001). To better understand how divergent selection operating between the sexes drives dimorphism, we require holistic approaches that fully describe the traits of both males and females (Riebel et al., 2019). Ideally, studies should include multiple signaling traits that mediate inter- and intra-sexual interactions (Bro-Jørgensen, 2010; Hebets et al., 2016; Hebets & Papaj, 2004; Partan & Marler, 2005), as well as their fitness consequences.

Bird song and plumage evolution offer excellent examples for how female, as well as male, signal evolution can lead to dimorphism. Females sang in the ancestor of modern songbirds and still sing in many tropical and temperate oscines (László Zsolt Garamszegi et al., 2006; Odom et al., 2014; Price et al., 2009), suggesting that losses of female song have most often created sex-based differences in song. Similarly, in species in which females have dull plumage, there has been repeated, independent selection on feather color in both males and females (Dale et al., 2015; Hofmann et al., 2008; Price et al., 2014), with females experiencing more evolutionary plumage change than males in some lineages (Price et al., 2014). Even though losses of female song and plumage appear to contribute considerably to dimorphism in signaling traits, the processes and conditions leading to these changes within individual species have not received much attention, particularly in species where female song is rare (Brunton et al., 2016; Kleindorfer et al., 2016a).

Large-scale evolutionary change in dimorphic signalling traits is generally assumed to result because some signals increase fitness. Numerous studies have identified male plumage traits associated with strong female preferences and increased fitness (Byers et al., 2010; Møller, 1988; Rodríguez et al., 2006; Ryan et al., 2019; Searcy & Andersson, 1986). In contrast, relatively few studies have found relationships between female visual signals and fitness proxies, but see: (Amundsen et al., 1997; Bulluck et al., 2017; Cain & Ketterson, 2011; Jawor et al., 2004; Pap et al., 2019). Likewise, many studies have examined the fitness correlates of male song (Catchpole & Slater, 2003; Gil & Gahr, 2002), but only a handful have looked at how female song impacts fitness (Brunton et al., 2016; Cain et al., 2015; Kleindorfer et al., 2016b; Krieg & Getty, 2016a). Although integrated approaches can best untangle the relative contributions of selection on males and females to these signaling modalities (Hebets et al., 2016; Riebel et al., 2019), researchers have seldom included both visual and acoustic phenotypes from both sexes within the same studies. Moreover, to our knowledge, no study has directly tested whether the pattern of dimorphism in vocal and visualsexual signaling traits is better predicted by contemporary selection on males (as is often assumed) or females. This likely reflects the difficulty of collecting reproductive success in the field, a lack of study in females, and a resulting paucity of datasets where the same traits and associated reproductive success are measured for both sexes. This is a particular challenge in species with infrequent female song, although such species provide a fertile testing ground for asking questions about why female song is reduced and differs from male song.

We propose that North American barn swallows (*Hirundo rustica*) offer a valuable system in which to investigate the selection pressures governing elaborate trait evolution and dimorphism. We explored patterns of plumage and song in a population of male and female barn swallows by: 1) Describing female song in this species, 2) exploring the form and seasonal timing of song production in both sexes, 3) determining which acoustic and visual traits are robustly dimorphic, and 4) testing whether sexual dimorphism is more strongly associated with

contemporary selection (i.e. trait correlations with reproductive success) in males or females. Much previous research is predicated on the idea that dimorphism results from selection for male trait exaggeration. Here, we test that assumption by including parallel data for both sexes and testing for linear associations between a measure of current selection and the magnitude of trait sexual dimorphism (similar to A. V. Badyaev & Martin, 2000). If dimorphism is driven/maintained by selection on males (for showier or elaborate traits), we expect that trait correlations with male reproductive success will predict degree of dimorphism. However, if dimorphism is driven/maintained by selection on females (for more cryptic or energy-efficient communication), we expect that trait associations with female reproductive success will predict dimorphism.

METHODS:

Background

As the subject of hundreds of studies on sexual signal evolution since the late 1980s, barn swallows offer an excellent system for testing questions of sexual dimorphism. Barn swallows are a weakly dimorphic species comprising six Holarctic subspecies that have rapidly diversified from a common ancestor as recently as 7,700 years ago (Smith et al., 2018). In North American barn swallows (*H. r. erythrogaster*), individuals with darker ventral feathers have higher reproductive success, while tail feather length is not a preferred trait in males (Rebecca J. Safran et al., 2016; R. J. Safran, 2004), or females (R. J. Safran, 2004). Previous barn swallow studies indicate that different components of male song (e.g. song rate and duration, rattle frequency and length) may provide receivers with information on the age, viability, condition, motivation and/or overall quality of singers (Dreiss et al., 2008; Galeotti et al., 1997, 2001; László Z. Garamszegi et al., 2005; L. Z. Garamszegi et al., 2006; Saino et al., 2003; Wilkins et al., 2015). Despite intensive study of this species, female signals have received little attention. While there is a modest body of research into male song, there are no indexed papers describing female barn swallow vocalizations in any detail. Females have been reported to produce a “twitter-warble” song (Brown & Brown, 2019), but no formal descriptions or quantitative analyses of that song exist (See Supplementary Methods and Background for clarifying notes on female song).

General Field Methods

Barn swallows in this study were part of a long-term study population across multiple sites in Boulder County, Colorado, ranging from 8 to 64 banded individuals. Persistent mist netting effort throughout the breeding season resulted in the capture and unique marking of every or nearly every bird present at each study site (using color bands, and application of Sharpie marker ink combinations to left and right rectrices). At the time of capture, tail streamer length was measured and a small sample (<90µL) of blood was taken and stored in 2% lysis buffer for later parentage analysis. Additionally, a set of ~5 contour feathers were sampled along a ventral transect (throat, breast, belly, and vent), and attached to index cards with tape for color analysis using a spectrophotometer in the lab (Hubbard et al., 2017; Wilkins et al., 2015). Nests were closely monitored for egg-laying activity, with checks conducted at least every four days to determine clutch initiation date, and date of hatching. On approximately day 12 post-hatching, all nestlings in a nest were banded and blood samples taken for parentage analysis.

All field methods were approved by the University of Colorado Institutional Animal Care and Use Committee (Protocols 07-07-SAF-01 and 1004.01).

Visual Trait Measures

We measured a total of 12 traits to characterize sexual signaling color variation: three axes of color (average brightness, hue, and red chroma) along a four-patch ventral transect (throat,

breast, belly, vent). We also measured maximum (rather than mean, as it is more normally distributed) tail streamer length (See Table S1 for trait definitions). Color and tail length were chosen for inclusion because they are known to have signal value in this species (Rebecca J. Safran et al., 2016; R. J. Safran, 2004; R. J. Safran et al., 2005). See Wilkins (2015) for detailed color methods.

Fitness Measures

We used female seasonal fledging success and male within-pair genetic paternity as fitness proxies for males and females, respectively. Female fledging success was calculated as the number of offspring leaving the nest across all broods. Genetic paternity was calculated as the number of genetic offspring within the social nest across all broods, determined through genetic paternity exclusions. Briefly, genotypes were derived from fluorescent-labeled PCR products of six microsatellite loci. Paternity was conducted using CERVUS v3.03 (Kalinowski et al., 2007) and an offspring was considered extra-pair when the offspring-mother-father trio confidence did not reach the 95% level. Detailed paternity exclusion methods are in Wilkins (2015). Four of the 12 fully-sampled males and females were social pairs, though we did not treat this statistically as the fitness correlational analysis was conducted separately for each sex.

Recordings

Female barn swallows were recorded as part of a study of male song (Wilkins et al., 2015) conducted in 2011 and 2012. Though males were the target of the prior study, when a female vocalization was heard, Author 1 immediately redirected the microphone at the source of the sound, capturing as many vocalizations as possible. Our estimates of relative singing outputs should therefore reflect realistic activity patterns. Songs were recorded in 16-bit WAV format, with 48kHz sampling rate using a Marantz PMD 660 paired with a Sennheiser MKH 20 and Telinga parabola (2011), or a Marantz PMD 661 paired with a Sennheiser ME62/k6 microphone and Telinga parabola (2012). Individuals were identified by a unique combination of colored ink applied to rectrices and colored leg bands applied during banding at the beginning of the field season. Total recording time was approximately 57 hours (May 2-May 31) in 2011 and 48 hours (May 1-Aug 21) in 2012, all between 5am and noon (Wilkins et al., 2015).

Acoustic Trait Measures

Using Raven v1.5 (Bioacoustics Research Program, 2011), we extracted acoustic parameters of all recorded songs from the 15 females (n=64) and a comparable number of songs (n=66) for 15 random males to create two acoustic datasets: (1) an element-level dataset of acoustic parameters measured for every element in all songs, and (2) a song-level dataset of acoustic parameters averaged for all elements in a song or calculated for the whole song. The element-level dataset was used to estimate an element diversity score, which estimates the variability, or diversity, of elements for a given song within an acoustic feature space of all barn swallow element measurements. This value was then appended to the song-level dataset, which was used for all subsequent statistical analyses involving song.

Of the 18 song-level variables extracted (see Supplementary Methods), we selected 11 that were likely to be robust to noise in recordings from barn swallow habitats and relevant for sexual signaling, based on previous studies of male song (Galeotti et al., 1997; L. Z. Garamszegi et al., 2006; Wilkins et al., 2015). The 11 selected acoustic traits were: song entropy (a measure of tonality), dominant frequency range, mean peak frequency, number of elements per song, element duration, element rate, song duration, gap duration, frequency interquartile range, and element diversity (see Table S1 for definitions). We then calculated individual means and coefficients of variation for each trait. This resulted in 22 traits describing differences in the magnitude and variability of different aspects of song frequency, timing, and

elemental diversity for both sexes. Song traits were further reduced to avoid redundancy in analysis of dimorphism and selection, as described below.

Statistical Analysis

Our sample included 15 females with at least one song measured. To make female-male comparisons that are not skewed by uneven sample sizes, we randomly selected 15 males with high quality recordings from our data set for paired analysis. For both sexes, due to missing data, we had 12 individuals with acoustic, visual, and fitness data. Because three females only had a single song recorded and coefficient of variation could not be calculated, only nine females had data for all 36 variables (22 acoustic traits,13 visual traits, and a fitness proxy). Thus, correlations between a particular trait and fitness metric in females were calculated from sample sizes ranging from nine to 15 individuals. Although our sample sizes are limited, the data set represents extensive population sampling time, and provides a starting point and methodology for more concerted effort..

All analyses were conducted using R v3.6.0 (R Core Team, 2019). To measure the magnitude of sexual dimorphism for each trait, we used Cohen's d, calculated as the difference of male from female population means divided by the pooled standard deviation (Cohen, 1988, p. 21). Thus, negative dimorphism values indicate male-biased dimorphism and positive values indicate female-biased dimorphism. For example, in barn swallows, tail feathers show male-biased dimorphism (as males have longer tails than females on average), resulting in a negative value on our dimorphism scale. To determine which traits were robustly dimorphic, we created 10,000 bootstraps using the 'bootstraps' function in the {rsample} package v0.0.5 (Kuhn et al., 2019), resampling sets of 15 males and 15 females with replacement. We then calculated 95% confidence intervals from this posterior distribution using the base function 'quantile.'

To avoid collinearity in further analysis of dimorphic traits, we identified pairs of dimorphic traits with Spearman's rank correlations >|0.7|, and retained the most biologically intuitive trait for each pair. For example, average number of elements, song duration, and our measure of element diversity showed pairwise correlations >0.8. As element diversity is a higher level measure of complexity, for which there is a wide literature suggesting it to be a common target of selection (Benedict & Najar, 2019; Snyder & Creanza, 2019), we retained mean element diversity in our set of dimorphic traits, and discarded mean song duration and mean number of elements. After filtering out redundant and non-dimorphic traits, we were left with 10 visual and acoustic traits for exploring the connection between dimorphism and contemporary selection pressures.

Testing trait-wise associations between reproductive success and dimorphism

As a measure of contemporary selection pressures, we calculated each trait's Spearman's rank correlation with a fitness surrogate (seasonal number of fledged offspring for females and seasonal number of within-pair genetic offspring for males). Finally, in order to test the linear relationship between measures of selection and observed trait dimorphism, while accounting for nonindependence of traits, we used the 'crunch' function in the {caper} package v1.0.1 (Orme et al., 2018), which implements the CAIC (Comparative Analysis by Independent Contrasts) algorithm (Purvis & Rambaut, 1995). Following a similar approach to Garamszegi et al. (2006), we treated traits as species with known phylogenetic relationships (i.e. correlation structure). To build the tree used by 'crunch,' we calculated trait distances as 1 - abs(Spearman's correlation), and constructed the tree using a single-linkage clustering method with the base 'hclust' function.

Thus, we made a linear model with dimorphism as a response and a trait's correlation with reproductive success as a predictor, controlling for correlation structure among traits. Given our

metric of dimorphism, we would expect a significant positive association between a trait's correlation with reproductive success and the magnitude of dimorphism if female fitness is maintaining sex-based trait differences. That is, traits that confer fitness benefits to females should have high dimorphism for our metric, and traits that reduce fecundity in females should have low dimorphism values. We would expect the opposite (a negative slope) if male fitness is driving/maintaining dimorphism. That is, traits associated with higher paternity would be more exaggerated in males (lower dimorphism values on our metric), and traits that reduce paternity would be less exaggerated in males (higher dimorphism values), resulting in a negative slope.

RESULTS

Characterizing female song

Recordings and observations revealed that females used a distinctive vocalization in response to hearing other females, in a manner synonymous to male counter-singing. They also were observed using the vocalization to interrupt the songs of their mates (see supplementary video 1). This vocalization is relatively complex, including many syllables with similar frequency modulation and pace to male song (Figure 1). Collectively, this evidence confirms that female barn swallows do produce facultative song, rather than just simple calls (Langmore, 1998).

Female song bouts were usually short and infrequent. As such, only 78 songs were recorded from 18 identifiable females over the course of two years, compared to 753 (865% more) songs from 40 males, given the same recording effort. This amounts to a recording efficiency of about 7.17 songs/hr for males and 0.74 songs/hr for females over the course of the study. The sexes also differed strikingly in the phenology of song production. Male songs were observed and recorded between May 1st and August 21st, while female songs were only observed and recorded between May 10th and May 29th (pooling years) (Figure 2a). For males, these recording dates translate to 59 days before to 109 days after first clutch initiation (i.e. breeding onset) by a males' mate with an average date (weighted by number songs recorded each day) of 5.43 days *after* breeding onset; Figure 2b). For females the range was 29 days before to 14 days after breeding onset, with a weighted average of 3.90 days *before* breeding onset. Due to dropouts due to recording quality, the total number of songs analyzed was 48 for females (mean=4.0, range= [1,11], sd=3.02) and 66 for males (mean=4.4, range= [2,5], sd= 0.99).

Morphological and song dimorphism in males and females

Our bootstrap analysis of 35 acoustic and visual traits showed that 16 were robustly dimorphic– i.e. had bootstrap 95% confidence intervals that did not overlap zero (Figure 3). These include a mixture of song and visual traits, with song traits being most distinctive between males and females. Most notably, female songs are brief in duration, include few elements, include elements with high entropy (i.e. noise), and have low element diversity across songs (Figure 3). Trait means, sample sizes, and dimorphism confidence intervals are shown in Table S2.

Connecting contemporary selection to trait dimorphism

Our final objective was to test whether sexual dimorphism is more strongly associated with contemporary selection in males or females. After controlling for intercorrelations between the 10 nonredundant, dimorphic traits using independent contrasts, we found a significant positive relationship between trait dimorphism and current selection pressures (i.e. correlation with a sex-appropriate fitness surrogate) in females (df=8, slope=6.04, t=2.92, p=0.019), but not males (df=8, slope=1.06, t=0.547, p=0.599) (Figure 4). That is, traits that had a negative association with female reproductive success (see Figure S1) showed male-biased sexual dimorphism (i.e., greater trait expression in males), while traits that positively correlated with female reproductive success showed female-biased sexual dimorphism. The relationship remained significant for females when including all 16 robustly dimorphic traits (df=14, slope=4.93, t=3.53, p=0.003), as

well as for the subset of acoustic dimorphic traits (df=11, slope=5.28, t=3.70, p=0.003). In contrast, the relationship between male traits and within-pair genetic paternity did not predict level of dimorphism for either of these subsets (all p≥0.260).

DISCUSSION:

Despite intensive study of this species for decades, female song in barn swallows has been reported as absent by some authors, the same as male song by others, or described as a solicitation call (see Supplementary Methods for notes on previous descriptions). Our data confirm that female barn swallows sing. We report that while female song is nearly nine times less frequent than male song, and relegated largely to the first month of the breeding season, it has an overall similar structure to male song and is used similarly in counter-singing, as well as mate-interruption contexts. The use of female song in counter-singing (see supplementary video 1) just prior to clutch initiation is consistent with singing patterns of other Northern temperate breeding species and suggests a functional role for this signal similar to what has been found for female song in those species (Cain et al., 2015; Krieg & Getty, 2016b; Levin, 1996; Rose et al., 2019; Yasukawa, 1989, 1990). Specifically, female barn swallow song could be actively maintained and used by females during this short period of the breeding season when females are establishing and competing for nest sites (Krieg & Getty, 2016b; Rosvall, 2011). Further, the use of female song to interrupt male song fits with theories that describe song functionality in mate attraction, pair bonding, and mate signal jamming (Grafe et al., 2004; Tobias & Seddon, 2009).

Although female and male barn swallow songs have a generally similar form, our results indicate that female barn swallow songs are shorter, with fewer elements, reduced frequency ranges, and lower element diversity (Figures 3, 4, S1). In this species, song traits are more strongly dimorphic (13 of 22 measured) than plumage traits (3 of 13 measured). This is consistent with the theory that elaborate female song may be selected against in certain species (Kleindorfer et al., 2016a; Wallace, 1868), and provides a mechanism supporting the loss of song in many North Temperate female birds (Odom et al., 2014). This is significant because, while the vast literature on male birdsong has informed many aspects of sexual selection theory (Gil & Gahr, 2002; Nowicki et al., 2002), and male song has been shown to play a key role in premating isolation and speciation (Alcaide et al., 2014; Irwin et al., 2001; Sosa-López et al., 2016; Wilkins et al., 2013, 2018), we know relatively little about female vocal signaling, even in classic study systems like barn swallows.

In contrast to a common assumption in the sexual selection literature, we found that overall patterns of dimorphism are better explained by selection on females, rather than on more elaborate males. That is, the overall magnitudes of sex-based differences in song form, tail streamer length, and color were predicted by how those traits correlated with female, but not male reproductive performance. Although our small sample size limits our ability to rigorously estimate selection gradients or account for error in our estimate of the correlation between traits and fitness proxies, these findings offer baseline evidence that selection on females may be more consistent and therefore more important in maintaining phenotypic distinctions between the sexes within populations. This is perhaps unsurprising, given previous studies showing variability in female preferences for male traits (Chaine & Lyon, 2008; Kingsolver et al., 2012; Robinson et al., 2008; Steele et al., 2011). Nevertheless, it has important implications for comparative studies which rely on the degree of plumage dimorphism as a surrogate for the strength of sexual selection. It is widely recognized that there are limitations to using dimorphism as a surrogate for sexual selection (Bell & Zamudio, 2012b; Huang & Rabosky, 2014; Kraaijeveld et al., 2011; Price, 2015b). We would add the caveat that researchers must consider the action of sexual selection on both males and females, and include the traits of both

sexes, even if traits are reduced or difficult to sample. Outcomes of this study suggest that following initial evolution of a mutually ornamented ancestor, dimorphism in a broad suite of visual and acoustic traits was most likely created and maintained in barn swallows through countervailing selection on females, rather than on males (Kraaijeveld et al., 2007).

Collectively, our results highlight the importance of studying rare behavioral phenomena and, in particular, call for better documentation and targeted research on female song (Odom & Benedict, 2018). We hope that future focused study of female songs will improve our understanding of the proximate causes and ultimate implications of these vocalizations. There are growing accounts of temperate-zone female birds that sing for short time periods during the breeding season, particularly during settlement of breeding territories (Hathcock & Benedict, 2018; Krieg & Getty, 2016c; Taff et al., 2012). Thus, more targeted research effort during this time period may very well show that complete loss of song in female Passerines is the exception, rather than the widely purported rule, and that similar selective pressures shape female song in multiple species. Our results also have implications for comparative studies which examine sexual dimorphism. With accumulating evidence for mutual mate choice and mutual ornamentation as an ancestral state across diverse taxa (Edward & Chapman, 2011; Kraaijeveld et al., 2007; Odom et al., 2014; Price et al., 2014), we offer a method for testing whether selection on females may contribute to dimorphism to a much greater extent than is currently appreciated. We believe that future work considering the evolution of both male and female communication systems in concert will greatly enhance our understanding of phenotypic variation and diversification.

Author Contributions: MRW collected data, performed analysis, and wrote initial manuscript draft; KJO measured and extracted acoustic metrics from audio files; LB undertook major restructuring of the manuscript; and RJS provided project administration and resources for carrying out the work. All authors contributed significantly to manuscript writing and revision.

Acknowledgements: MRW thanks numerous undergraduate and graduate students in the Safran Lab who helped collect the mountain of phenotypic and fitness data that enabled this study–and the many site owners who allowed us to traipse around their property at odd hours of the day (in particular, the Colorado Horse Rescue). We also thank Megan Barkdull for help with sound analysis.

Funding: This work was supported by an American Ornithological Society: Joseph Grinnell Award and an NSF Graduate Research Fellowship to MRW. KJO was supported by a U.S. National Science Foundation Postdoctoral Research Fellowship in Biology (grant no. 1612861).

References

Alcaide, M., Scordato, E. S. C., Price, T. D., & Irwin, D. E. (2014). Genomic divergence in a ring species complex. Nature, 511(7507), 83–85.

Amundsen, T., Forsgren, E., & Hansen, L. T. T. (1997). On the function of female ornaments: male bluethroats prefer colourful females. Proceedings of the Royal Society B: Biological Sciences, 264(1388), 1579–1586.

Andersson, M. B. (1994). Sexual Selection. Princeton University Press.

Badyaev, A. V., & Hill, G. E. (2003). Avian Sexual Dichromatism in Relation to Phylogeny and Ecology. In Annual Review of Ecology, Evolution, and Systematics (Vol. 34, Issue 1, pp. 27–49). https://doi.org/10.1146/annurev.ecolsys.34.011802.132441

Badyaev, A. V., & Martin, T. E. (2000). Sexual dimorphism in relation to current selection in the house finch. Evolution; International Journal of Organic Evolution, 54(3), 987–997.

Bell, R. C., & Zamudio, K. R. (2012a). Sexual dichromatism in frogs: natural selection, sexual selection and unexpected diversity. Proceedings. Biological Sciences / The Royal Society, 279(1748), 4687–4693.

Bell, R. C., & Zamudio, K. R. (2012b). Sexual dichromatism in frogs: natural selection, sexual selection and unexpected diversity. Proceedings. Biological Sciences / The Royal Society, 279(1748), 4687–4693.

Benedict, L., & Najar, N. A. (2019). Are commonly used metrics of bird song complexity concordant? The Auk, 136(1), 201.

Bioacoustics Research Program. (2011). Raven Pro: Interactive Sound Analysis Software (Version 1.5) [Computer software]. http://www.birds.cornell.edu/raven

Bro-Jørgensen, J. (2010). Dynamics of multiple signalling systems: animal communication in a world in flux. Trends in Ecology & Evolution, 25(5), 292–300.

Brown, M. B., & Brown, C. R. (2019). Barn Swallow (Hirundo rustica). The Birds of North America Online. https://doi.org/10.2173/bna.barswa.02

Brunton, D. H., Roper, M. M., & Harmer, A. M. T. (2016). Female Song Rate and Structure Predict Reproductive Success in a Socially Monogamous Bird. Frontiers in Ecology and Evolution, 4. https://doi.org/10.3389/fevo.2016.00013

Bulluck, L. P., Foster, M. J., Kay, S., Cox, D. E., Viverette, C., & Huber, S. (2017). Feather carotenoid content is correlated with reproductive success and provisioning rate in female Prothonotary Warblers. The Auk, 134(1), 229–239.

Burns, K. J. (1998). A Phylogenetic Perspective on the Evolution of Sexual Dichromatism in Tanagers (Thraupidae): The Role of Female versus Male Plumage. Evolution; International Journal of Organic Evolution, 52(4), 1219.

Byers, J., Hebets, E., & Podos, J. (2010). Female mate choice based upon male motor performance. In Animal Behaviour (Vol. 79, Issue 4, pp. 771–778). https://doi.org/10.1016/j.anbehav.2010.01.009

Cain, K. E., Cockburn, A., & Langmore, N. E. (2015). Female song rates in response to simulated intruder are positively related to reproductive success. Frontiers in Ecology and Evolution, 3. https://doi.org/10.3389/fevo.2015.00119

Cain, K. E., & Ketterson, E. D. (2011). Competitive females are successful females; phenotype, mechanism, and selection in a common songbird. Behavioral Ecology and Sociobiology, 66(2), 241–252.

Catchpole, C. K., & Slater, P. J. B. (2003). Bird Song: Biological Themes and Variations. Cambridge University Press.

Chaine, A. S., & Lyon, B. E. (2008). Adaptive plasticity in female mate choice dampens sexual selection on male ornaments in the lark bunting. Science, 319(5862), 459–462.

Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences (2nd ed.). LAWRENCE ERLBAUM ASSOCIATES.

Dale, J., Dey, C. J., Delhey, K., Kempenaers, B., & Valcu, M. (2015). The effects of life history and sexual selection on male and female plumage colouration. Nature, 527(7578), 367–370.

Darwin, C. (1859). On the origin of species by means of natural selection, or, The preservation of favoured races in the struggle for life /. https://doi.org/10.5962/bhl.title.68064

Darwin, E. (1801). Zoonomia Or the Laws of Organic Life.

Dreiss, A. N., Navarro, C., de Lope, F., & Møller, A. P. (2008). Effects of an Immune Challenge on Multiple Components of Song Display in Barn Swallows Hirundo rustica : Implications for Sexual Selection. Ethology: Formerly Zeitschrift Fur Tierpsychologie, 114(10), 955–964.

Edward, D. A., & Chapman, T. (2011). The evolution and significance of male mate choice. In Trends in Ecology & Evolution (Vol. 26, Issue 12, pp. 647–654). https://doi.org/10.1016/j.tree.2011.07.012

Evans, M. R. (1998). Selection on swallow tail streamers. In Nature (Vol. 394, Issue 6690, pp. 233–234). https://doi.org/10.1038/28297

Freed, L. A. (2000). Female ornamentation, mate choice and sexual selection. Trends in Ecology & Evolution, 15(11), 471.

Galeotti, P., Saino, N., Perani, E., Sacchi, R., & Møller, A. R. (2001). Age□related song variation in male barn swallows. In Italian Journal of Zoology (Vol. 68, Issue 4, pp. 305–310). https://doi.org/10.1080/11250000109356423

Galeotti, P., Saino, N., Sacchi, R., & Møller, A. P. (1997). Song correlates with social context, testosterone and body condition in male barn swallows. Animal Behaviour, 53(4), 687–700.

Garamszegi, L. Z., Hegyi, G., Heylen, D., Ninni, P., de Lope, F., Eens, M., & Møller, A. P. (2006). The design of complex sexual traits in male barn swallows: associations between signal attributes. Journal of Evolutionary Biology, 19(6), 2052–2066.

Garamszegi, L. Z., Heylen, D., Møller, A. P., Eens, M., & de Lope, F. (2005). Age-dependent health status and song characteristics in the barn swallow. Behavioral Ecology: Official Journal of the International Society for Behavioral Ecology, 16(3), 580–591.

Garamszegi, L. Z., Pavlova, D. Z., Eens, M., & Møller, A. P. (2006). The evolution of song in female birds in Europe. Behavioral Ecology: Official Journal of the International Society for Behavioral Ecology, 18(1), 86–96.

Gil, D., & Gahr, M. (2002). The honesty of bird song: multiple constraints for multiple traits. Trends in Ecology & Evolution, 17(3), 133–141.

Grafe, T. U., Ulmar Grafe, T., & Bitz, J. H. (2004). Functions of duetting in the tropical boubou, Laniarius aethiopicus: territorial defence and mutual mate guarding. In Animal Behaviour (Vol. 68, Issue 1, pp. 193–201). https://doi.org/10.1016/j.anbehav.2003.11.003

Hathcock, T. J., & Benedict, L. (2018). Conspecific challenges provoke female canyon wrens to sing but not to duet. In Behavioral Ecology and Sociobiology (Vol. 72, Issue 12). https://doi.org/10.1007/s00265-018-2625-0

Hebets, E. A., Barron, A. B., Balakrishnan, C. N., Hauber, M. E., Mason, P. H., & Hoke, K. L. (2016). A systems approach to animal communication. Proceedings. Biological Sciences / The Royal Society, 283(1826), 20152889.

Hebets, E. A., & Papaj, D. R. (2004). Complex signal function: developing a framework of testable hypotheses. Behavioral Ecology and Sociobiology, 57(3), 197–214.

Hedrick, A. V., & Temeles, E. J. (1989). The evolution of sexual dimorphism in animals: Hypotheses and tests. Trends in Ecology & Evolution, 4(5), 136–138.

Hofmann, C. M., Cronin, T. W., & Omland, K. E. (2008). EVOLUTION OF SEXUAL DICHROMATISM. 1. CONVERGENT LOSSES OF ELABORATE FEMALE COLORATION IN NEW WORLD ORIOLES (ICTERUSSPP.). In The Auk (Vol. 125, Issue 4, pp. 778–789). https://doi.org/10.1525/auk.2008.07112

Huang, H., & Rabosky, D. L. (2014). Sexual selection and diversification: reexamining the correlation between dichromatism and speciation rate in birds. The American Naturalist, 184(5), E101–E114.

Hubbard, J. K., Hund, A. K., Levin, I. I., McGraw, K. J., Wilkins, M. R., & Safran, R. J. (2017). The importance of cross-validation, accuracy, and precision for measuring plumage color: A comment on Vaquero-Alba et al. (2016). The Auk, 134(1), 34–38.

Irwin, D. E., Bensch, S., & Price, T. D. (2001). Speciation in a ring. Nature, 409(6818), 333–337.

Jawor, J. M., Gray, N., Beall, S. M., & Breitwisch, R. (2004). Multiple ornaments correlate with aspects of condition and behaviour in female northern cardinals, Cardinalis cardinalis. Animal Behaviour, 67(5), 875–882.

Jennions, M. D., Kahn, A. T., Kelly, C. D., & Kokko, H. (2012). Meta-analysis and sexual selection: past studies and future possibilities. Evolutionary Ecology, 26(5), 1119–1151.

Kalinowski, S. T., Taper, M. L., & Marshall, T. C. (2007). Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. Molecular Ecology, 16(5), 1099–1106.

Kingsolver, J. G., Diamond, S. E., Siepielski, A. M., & Carlson, S. M. (2012). Synthetic analyses of phenotypic selection in natural populations: lessons, limitations and future directions. In Evolutionary Ecology (Vol. 26, Issue 5, pp. 1101–1118). https://doi.org/10.1007/s10682-012-9563-5

Kleindorfer, S., Evans, C., & Mahr, K. (2016a). Female in-nest chatter song increases predation. Biology Letters, 12(1), 20150513.

Kleindorfer, S., Evans, C., & Mahr, K. (2016b). Female in-nest chatter song increases predation. Biology Letters, 12(1), 20150513.

Kraaijeveld, K., Kraaijeveld-Smit, F. J. L., & Komdeur, J. (2007). The evolution of mutual ornamentation. In Animal Behaviour (Vol. 74, Issue 4, pp. 657–677). https://doi.org/10.1016/j.anbehav.2006.12.027

Kraaijeveld, K., Kraaijeveld-Smit, F. J. L., & Maan, M. E. (2011). Sexual selection and speciation: the comparative evidence revisited. Biological Reviews of the Cambridge Philosophical Society, 86(2), 367–377.

Krieg, C. A., & Getty, T. (2016a). Not just for males: females use song against male and female rivals in a temperate zone songbird. Animal Behaviour, 113, 39–47.

Krieg, C. A., & Getty, T. (2016b). Not just for males: females use song against male and female rivals in a temperate zone songbird. Animal Behaviour, 113, 39–47.

Krieg, C. A., & Getty, T. (2016c). Not just for males: females use song against male and female rivals in a temperate zone songbird. Animal Behaviour, 113, 39–47.

Kuhn, M., Chow, F., & Wickham, H. (2019). rsample: General Resampling Infrastructure (Version 0.0.5) [Computer software]. https://CRAN.R-project.org/package=rsample

Lande, R. (1980). SEXUAL DIMORPHISM, SEXUAL SELECTION, AND ADAPTATION IN POLYGENIC CHARACTERS. Evolution; International Journal of Organic Evolution, 34(2), 292–305.

Langmore, N. E. (1998). Functions of duet and solo songs of female birds. Trends in Ecology & Evolution, 13(4), 136–140.

Levin, R. N. (1996). Song behaviour and reproductive strategies in a duetting wren,Thryothorus nigricapillus: I. Removal experiments. In Animal Behaviour (Vol. 52, Issue 6, pp. 1093–1106). https://doi.org/10.1006/anbe.1996.0257

Miller, C. W., McDonald, G. C., & Moore, A. J. (2016). The tale of the shrinking weapon: seasonal changes in nutrition affect weapon size and sexual dimorphism, but not contemporary evolution. In Journal of Evolutionary Biology (Vol. 29, Issue 11, pp. 2266–2275). https://doi.org/10.1111/jeb.12954

Møller, A. P. (1988). Female choice selects for male sexual tail ornaments in the monogamous swallow. In Nature (Vol. 332, Issue 6165, pp. 640–642). https://doi.org/10.1038/332640a0

Ng, T. P. T., Rolán-Alvarez, E., Dahlén, S. S., Davies, M. S., Estévez, D., Stafford, R., & Williams, G. A. (2019). The causal relationship between sexual selection and sexual size dimorphism in marine gastropods. In Animal Behaviour (Vol. 148, pp. 53–62). https://doi.org/10.1016/j.anbehav.2018.12.005

Nowicki, S., Searcy, W. A., & Peters, S. (2002). Brain development, song learning and mate choice in birds: a review and experimental test of the “nutritional stress hypothesis.” Journal of Comparative Physiology. A, Neuroethology, Sensory, Neural, and Behavioral Physiology, 188(11-12), 1003–1014.

Odom, K. J., & Benedict, L. (2018). A call to document female bird songs: Applications for diverse fields. In The Auk (Vol. 135, Issue 2, pp. 314–325). https://doi.org/10.1642/auk-17-183.1

Odom, K. J., Hall, M. L., Riebel, K., Omland, K. E., & Langmore, N. E. (2014). Female song is widespread and ancestral in songbirds. Nature Communications, 5, 3379.

Orme, D., Freckleton, R., Thomas, G., Petzoldt, T., Fritz, S., Isaac, N., & and Pearse, W. (2018). caper: Comparative Analyses of Phylogenetics and Evolution in R (Version 1.0.1) [Computer software]. https://CRAN.R-project.org/package=caper

Pap, P. L., Fülöp, A., Adamkova, M., Cepak, J., Michalkova, R., Safran, R. J., Stermin, A. N., Tomasek, O., Vágási, C. I., Vincze, O., Wilkins, M. R., & Albrecht, T. (2019). Selection on multiple sexual signals in two Central and Eastern European populations of the barn swallow. In Ecology and Evolution. https://doi.org/10.1002/ece3.5629

Partan, S. R., & Marler, P. (2005). Issues in the classification of multimodal communication signals. The American Naturalist, 166(2), 231–245.

Price, J. J. (2015a). Rethinking our assumptions about the evolution of bird song and other sexually dimorphic signals. Frontiers in Ecology and Evolution, 3. https://doi.org/10.3389/fevo.2015.00040

Price, J. J. (2015b). Rethinking our assumptions about the evolution of bird song and other sexually dimorphic signals. In Frontiers in Ecology and Evolution (Vol. 3). https://doi.org/10.3389/fevo.2015.00040

Price, J. J., Jordan Price, J., & Eaton, M. D. (2014). RECONSTRUCTING THE EVOLUTION OF SEXUAL DICHROMATISM: CURRENT COLOR DIVERSITY DOES NOT REFLECT PAST RATES OF MALE AND FEMALE CHANGE. In Evolution (Vol. 68, Issue 7, pp. 2026–2037). https://doi.org/10.1111/evo.12417

Price, J. J., Lanyon, S. M., & Omland, K. E. (2009). Losses of female song with changes from tropical to temperate breeding in the New World blackbirds. Proceedings. Biological Sciences / The Royal Society, 276(1664), 1971–1980.

Purvis, A., & Rambaut, A. (1995). Comparative analysis by independent contrasts (CAIC): an Apple Macintosh application for analysing comparative data. Computer Applications in the Biosciences: CABIOS, 11(3), 247–251.

R Core Team. (2019). R: A Language and Environment for Statistical Computing (Version 3.6.0) [Computer software]. https://www.R-project.org/

Riebel, K. (2016). Understanding Sex Differences in Form and Function of Bird Song: The Importance of Studying Song Learning Processes. Frontiers in Ecology and Evolution, 4. https://doi.org/10.3389/fevo.2016.00062

Riebel, K., Hall, M. L., & Langmore, N. E. (2005). Female songbirds still struggling to be heard [Review of Female songbirds still struggling to be heard]. Trends in Ecology & Evolution, 20(8), 419–420.

Riebel, K., Odom, K. J., Langmore, N. E., & Hall, M. L. (2019). New insights from female bird song: towards an integrated approach to studying male and female communication roles. Biology Letters, 15(4), 20190059.

Robinson, M. R., Pilkington, J. G., Clutton-Brock, T. H., Pemberton, J. M., & Kruuk, L. E. B. (2008). Environmental heterogeneity generates fluctuating selection on a secondary sexual trait. Current Biology: CB, 18(10), 751–757.

Rodríguez, R. L., Ramaswamy, K., & Cocroft, R. B. (2006). Evidence that female preferences have shaped male signal evolution in a clade of specialized plant-feeding insects. Proceedings. Biological Sciences / The Royal Society, 273(1601), 2585–2593.

Rose, E. M., Coss, D. A., Haines, C. D., Danquah, S. A., Studds, C. E., & Omland, K. E. (2019). Why do females sing?—pair communication and other song functions in eastern bluebirds. In Behavioral Ecology (Vol. 30, Issue 6, pp. 1653–1661). https://doi.org/10.1093/beheco/arz130

Rosvall, K. A. (2011). Intrasexual competition in females: evidence for sexual selection? Behavioral Ecology: Official Journal of the International Society for Behavioral Ecology, 22(6), 1131–1140.

Ryan, M. J., Akre, K. L., Baugh, A. T., Bernal, X. E., Lea, A. M., Leslie, C., Still, M. B., Wylie, D. C., & Rand, A. S. (2019). Nineteen Years of Consistently Positive and Strong Female Mate Preferences despite Individual Variation. The American Naturalist, 194(2), 125–134.

Safran, R. J. (2004). Plumage coloration, not length or symmetry of tail-streamers, is a sexually selected trait in North American barn swallows. In Behavioral Ecology (Vol. 15, Issue 3, pp. 455–461). https://doi.org/10.1093/beheco/arh035

Safran, R. J., Neuman, C. R., McGraw, K. J., & Lovette, I. J. (2005). Dynamic paternity allocation as a function of male plumage color in barn swallows. Science, 309(5744), 2210–2212.

Safran, R. J., Vortman, Y., Jenkins, B. R., Hubbard, J. K., Wilkins, M. R., Bradley, R. J., & Lotem, A. (2016). The maintenance of phenotypic divergence through sexual selection: An experimental study in barn swallows Hirundo rustica. Evolution; International Journal of Organic Evolution, 70(9), 2074–2084.

Saino, N., Romano, M., Sacchi, R., Ninni, P., Galeotti, P., & Møller, A. P. (2003). Do male barn swallows (Hirundo rustica) experience a trade-off between the expression of multiple sexual signals? Behavioral Ecology and Sociobiology, 54(5), 465–471.

Searcy, W. A., & Andersson, M. (1986). Sexual Selection and the Evolution of Song. Annual Review of Ecology and Systematics, 17(1), 507–533.

Smith, C. C. R., Flaxman, S. M., Scordato, E. S. C., Kane, N. C., Hund, A. K., Sheta, B. M., & Safran, R. J. (2018). Demographic inference in barn swallows using whole-genome data shows signal for bottleneck and subspecies differentiation during the Holocene. Molecular Ecology, 27(21), 4200–4212.

Snyder, K. T., & Creanza, N. (2019). Polygyny is linked to accelerated birdsong evolution but not to larger song repertoires. Nature Communications, 10(1), 884.

Sosa-López, J. R., Martínez Gómez, J. E., & Mennill, D. J. (2016). Divergence in mating signals correlates with genetic distance and behavioural responses to playback. Journal of Evolutionary Biology, 29(2), 306–318.

Steele, D. B., Siepielski, A. M., & McPEEK, M. A. (2011). Sexual selection and temporal phenotypic variation in a damselfly population. In Journal of Evolutionary Biology (Vol. 24, Issue 7, pp. 1517–1532). https://doi.org/10.1111/j.1420-9101.2011.02284.x

Taff, C. C., Littrell, K. A., & Freeman-Gallant, C. R. (2012). Female Song in the Common Yellowthroat. In The Wilson Journal of Ornithology (Vol. 124, Issue 2, pp. 370–374). https://doi.org/10.1676/11-182.1

Tobias, J. A., & Seddon, N. (2009). Signal jamming mediates sexual conflict in a duetting bird. Current Biology: CB, 19(7), 577–582.

Wallace, A. R. (1868). A Theory of Birds' Nests; showing the relation of certain differences of colour in female birds to their mode of nidification. In Contributions to the Theory of Natural Selection (pp. 231–263). https://doi.org/10.1017/cbo9780511693106.009

Wiens, J. J. (2001). Widespread loss of sexually selected traits: how the peacock lost its spots. In Trends in Ecology & Evolution (Vol. 16, Issue 9, pp. 517–523). https://doi.org/10.1016/s0169-5347(01)02217-0

Wilkins, M. R., Scordato, E. S. C., Semenov, G. A., Karaardiç, H., Shizuka, D., Rubtsov, A., Pap, P. L., Shen, S.-F., & Safran, R. J. (2018). Global song divergence in barn swallows (Hirundo rustica): exploring the roles of genetic, geographical and climatic distance in sympatry and allopatry. In Biological Journal of the Linnean Society (Vol. 123, Issue 4, pp. 825–849). https://doi.org/10.1093/biolinnean/bly012

Wilkins, M. R., Seddon, N., & Safran, R. J. (2013). Evolutionary divergence in acoustic signals: causes and consequences. Trends in Ecology & Evolution, 28(3), 156–166.

Wilkins, M. R., Shizuka, D., Joseph, M. B., Hubbard, J. K., & Safran, R. J. (2015). Multimodal signalling in the North American barn swallow: a phenotype network approach. Proceedings. Biological Sciences / The Royal Society, 282(1816), 20151574.

Yasukawa, K. (1989). The costs and benefits of a vocal signal: the nest-associated "Chit" of the female red-winged blackbird, Agelaius phoeniceus. In Animal Behaviour (Vol. 38, Issue 5, pp. 866–874). https://doi.org/10.1016/s0003-3472(89)80118-6

Yasukawa, K. (1990). Does the "teer" vocalization deter prospecting female red-winged blackbirds? Behavioral Ecology and Sociobiology, 26(6), 421–426.

Figure 1: Spectrograms of female and male barn swallows. We show two song renditions for 3 females and 3 males to demonstrate variability. Photos by Author 1.

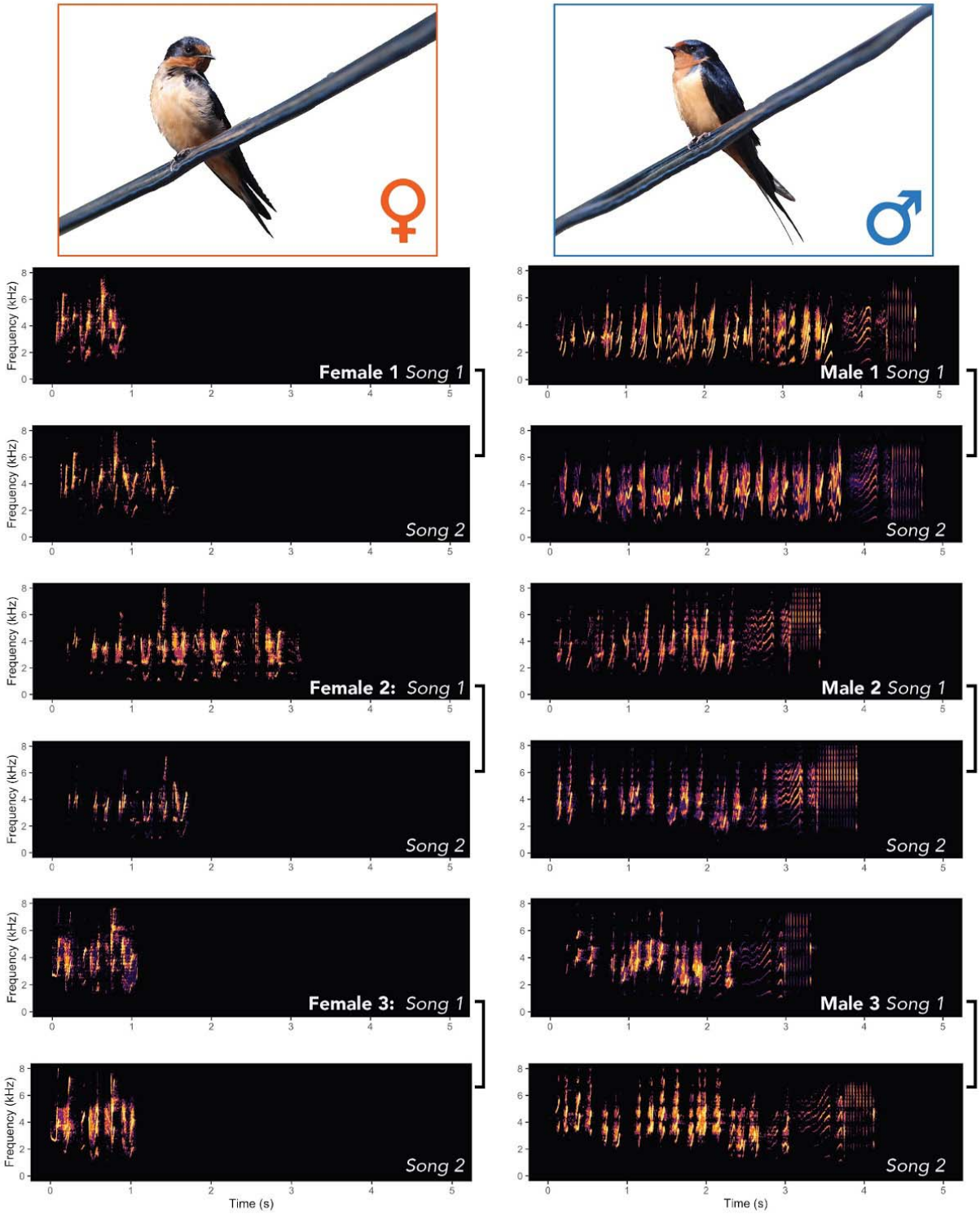


Figure 2: Seasonal variation in song production, as demonstrated by opportunistic recording of spontaneous male and female songs. A) Shows number of songs recorded for both sexes, as a function of recording date. In B) the day of breeding onset (first clutch initiation) has been subtracted from Julian recording day; thus, negative x-values represent recordings taken before breeding onset. The dashed vertical line marks Day 4 post clutch initiation–i.e. the last day of female fertility before laying of the ultimate egg of a modal 5-egg clutch. Lines connect points for individuals with recordings on multiple days.

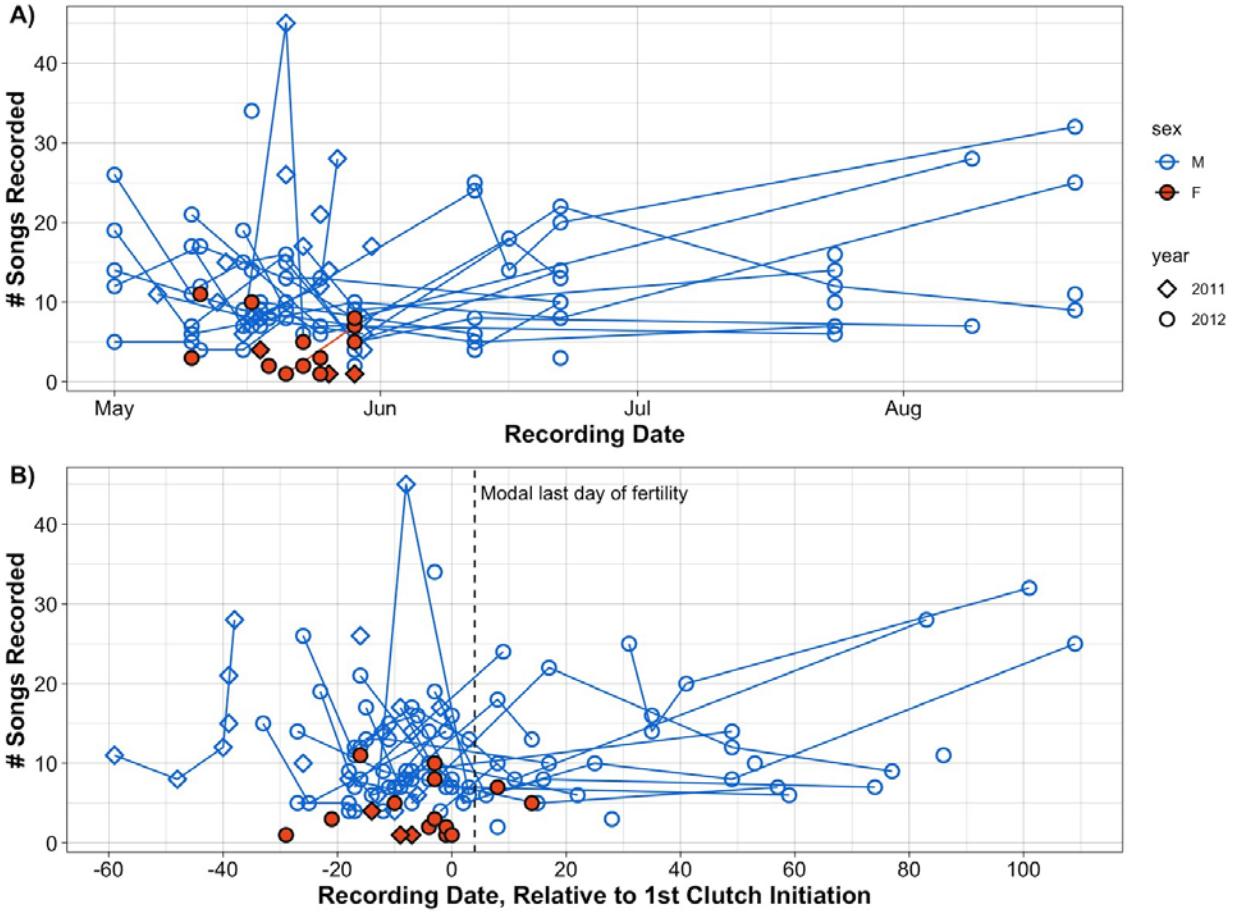


Figure 3: Sexual dimorphism for 35 potential sexual signaling traits. Empirical point estimates for Cohen's d are shown, with 95% bootstrapped confidence intervals not overlapping zero bolded in purple. Symbols for points indicate trait modality (circle=acoustic, triangle=visual). Traits are means, except where preceded by CV (coefficient of variation). Trait names beginning with a star were the chosen set of dimorphic traits for further analysis, after removing highly redundant traits with Spearman's correlations ≥|0.7|.

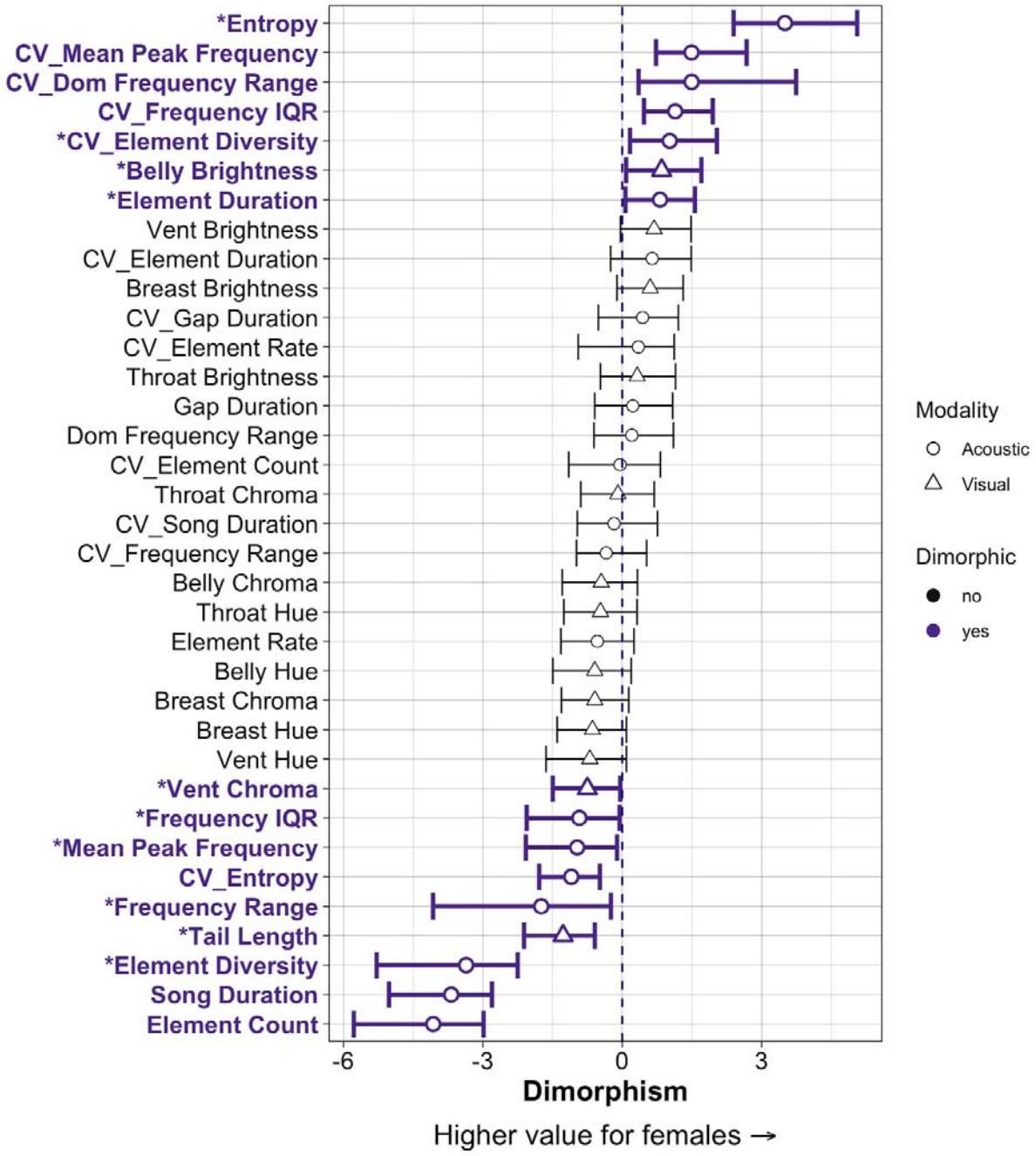
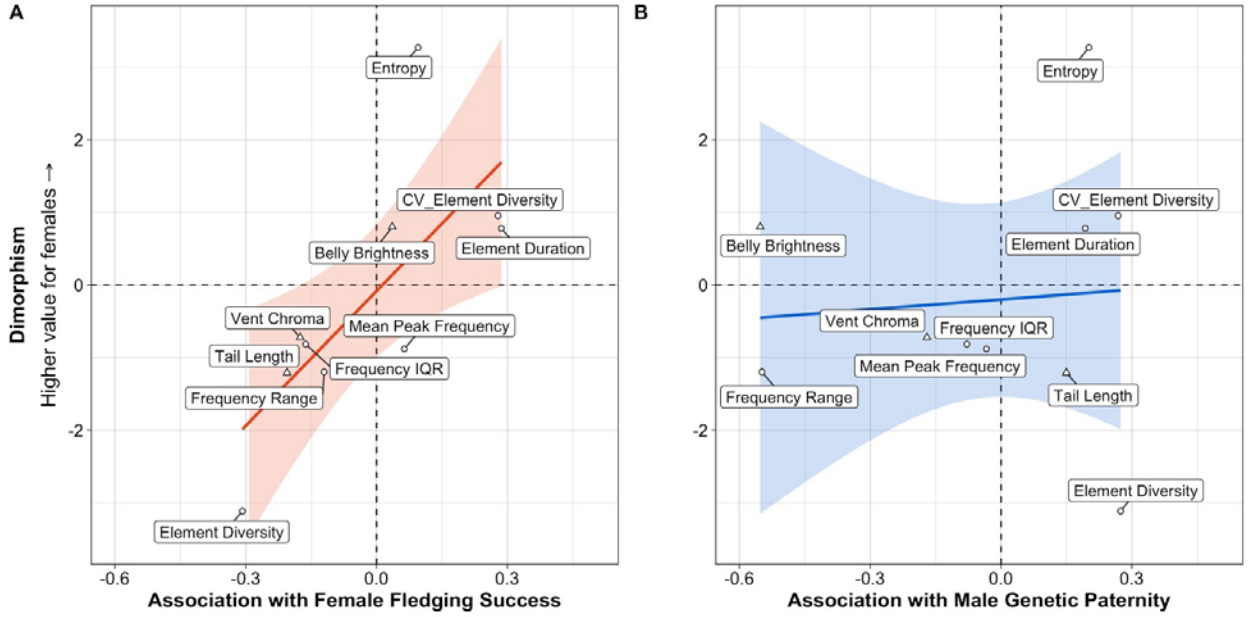


Figure 4: Relationship between trait dimorphism and trait association with reproductive output for A) females and B) males. Dimorphism was measured as the difference of scaled female and male trait values in pooled standard deviation units (Cohen's d). Reproductive output was measured as seasonal fledging success for females and within-pair seasonal genetic paternity for males. Females showed a much stronger (and significant) relationship between how traits impacted reproductive performance and the observed amount of sexual dimorphism.



Supplementary Figures and Tables

Figure S1. Spearman's rank correlations between dimorphic traits shown in Figure 4 and a relevant fitness metric–our measure of contemporary selection. The fitness metric for males was seasonal within pair genetic paternity, and for females it was seasonal fledging success. Orange 95% confidence ellipses and points are for females; blue ellipses and points are for males. Plots are ordered (top left to bottom right) from lowest to highest correlation for females.

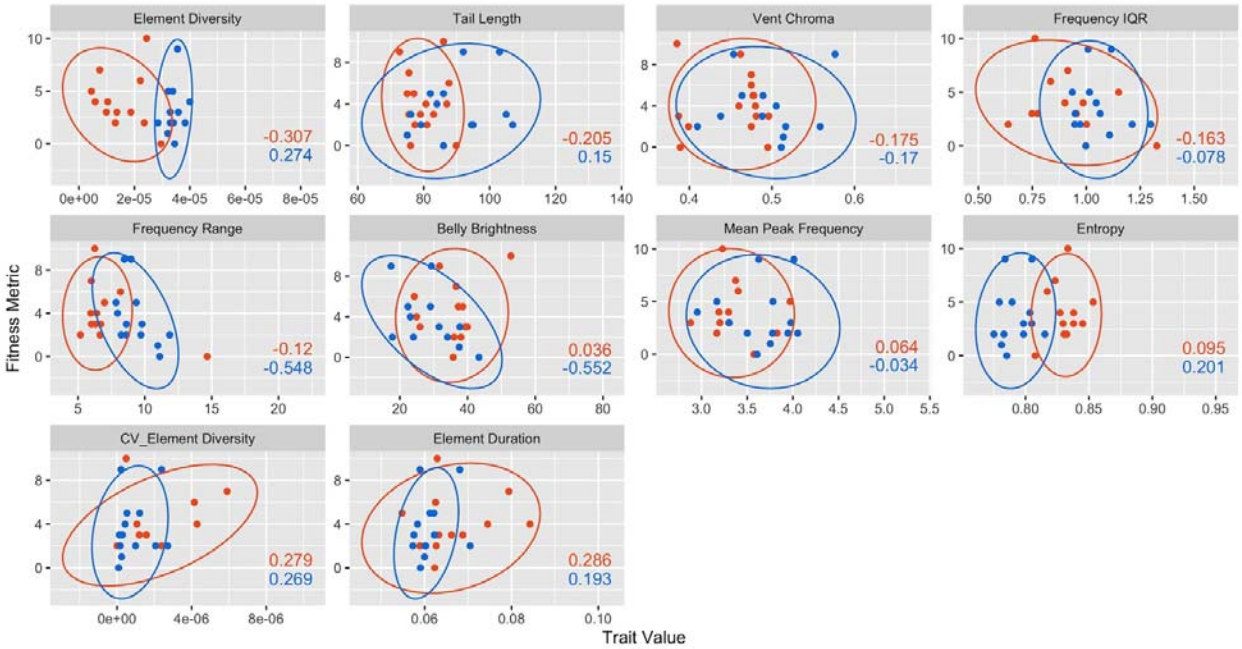


Table S1. Trait definitions for all per individual visual and acoustic variables calculated for this study.

Trait Type	Trait	Definition
Acoustic	Entropy	Average spectrographic entropy: the product of time and energy distribution of the frequency spectrum; unitless, ranging from 0 (pure tone) to 1 (white noise)
Acoustic	Dominant Frequency Range	Average range of the dominant frequency measured across the song
Acoustic	Frequency Range	Average range of raw frequencies measured across song
Acoustic	Mean Peak Frequency	Average frequency with highest energy from the mean frequency spectrum
Acoustic	Element Count	Number of song elements (syllables) per song
Acoustic	Element Duration	Average duration of elements per individual (in s)
Acoustic	Element Rate	Average rate of element production (number of elements divided by duration) per song (in elements/s)
Acoustic	Song Duration	Average duration of songs (in s)
Acoustic	Gap Duration	Average length of intervals between elements (in s)
Acoustic	Element Diversity	Average minimum convex polygon area surrounding the element-level acoustic space for each song
Acoustic	Frequency IQR	Average interquartile range for frequency (in kHz)
Acoustic	CV_Entropy	Coefficient of variation for spectrographic entropy: the product of time and frequency entropy
Acoustic	CV_Dom Frequency Range	Coefficient of variation for the range of the dominant frequency measured across the song
Acoustic	CV_Frequency Range	Coefficient of variation for raw frequencies measured across song
Acoustic	CV_Mean Peak Frequency	Coefficient of variation for frequency with highest energy from the mean frequency spectrum
Acoustic	CV_Element Count	Coefficient of variation for the number of song elements (syllables) per song
Acoustic	CV_Element Duration	Coefficient of variation for duration of elements (in s)
Acoustic	CV_Element Rate	Coefficient of variation for rate of element production per song (in elements/s)

Acoustic	CV_Song Duration	Coefficient of variation for duration of songs (in s)
Acoustic	CV_Gap Duration	Coefficient of variation for length of intervals between elements (in s)
Acoustic	CV_Element Diversity	
Acoustic	CV_Frequency IQR	Coefficient of variation for interquartile range for frequency (in kHz)
Visual	Tail Length	Maximum length of two longest tail feathers (streamers), each of which was the average of 3 repeated measures (in mm)
Visual	Brightness	The average percent reflectance between 300 and 700 nm; lower values are darker; measured for Throat, Breast, Belly, and Vent
Visual	Hue	The wavelength at the maximum slope of reflectance spectrum (in nm); low values pale/yellowish; high values dark/reddish; measured for Throat, Breast, Belly, and Vent
Visual	Chroma	The proportion of light reflected in the red color range (600-700nm); higher values darker; measured for Throat, Breast, Belly, and Vent

Table S2. Female and male, sample sizes, means, standard deviations, dimorphism (i.e. Cohen's D), and dimorphism 95% confidence intervals for 10,000 bootstraps. Sample sizes for coefficients of variation are lower for females, as three individuals only had a single song recorded. Bolded traits had dimorphism confidence intervals that did not overlap zero. Starred traits were selected for the final analysis in Figure 4, after eliminating the less biologically intuitive trait for redundant trait pairs with >|0.7| correlations.

Trait	F_n	F_Mean (SD)	M_n	M_mean (SD)	Dimorphism	95% Conf Int
***Belly Brightness**	**13**	**35.5 (7.62)**	**14**	**29.3 (7.62)**	**0.799**	**(0.0694, 1.31)**
Belly Chroma	13	0.421 (0.0434)	14	0.441 (0.0434)	-0.425	(-1.09, 0.333)
Belly Hue	13	608 (29)	14	626 (29)	-0.554	(-1.21, 0.19)
Breast Brightness	15	35.7 (9.39)	14	30.8 (9.39)	0.589	(-0.114, 1.1)
Breast Chroma	15	0.424 (0.0356)	14	0.446 (0.0356)	-0.577	(-1.1, 0.129)
Breast Hue	15	616 (23.3)	14	631 (23.3)	-0.62	(-1.16, 0.0884)
CV_Dom Frequency Range	**9**	**0.0902 (0.0483)**	**15**	**0.0363 (0.0483)**	**1.27**	**(0.362, 1.79)**
CV_Element Count	9	1.5 (2.24)	15	1.41 (2.24)	0.0468	(-0.962, 0.875)
***CV_Element Diversity**	**9**	**2.34e-06 (2.01e-06)**	**15**	**8.61e-07 (2.01e-06)**	**0.954**	**(0.181, 1.57)**
CV_Element Duration	9	0.000904 (0.000889)	15	0.000465 (0.000889)	0.626	(-0.264, 1.34)
CV_Element Rate	9	0.149 (0.188)	15	0.0899 (0.188)	0.432	(-0.903, 1.15)
CV_Entropy	**9**	**0.000127 (0.000117)**	**15**	**0.000359 (0.000117)**	**-1.02**	**(-1.29, -0.454)**
CV_Frequency IQR	**9**	**0.069 (0.066)**	**15**	**0.0189 (0.066)**	**1.04**	**(0.517, 1.58)**
CV_Frequency Range	9	0.25 (0.573)	15	0.497 (0.573)	-0.319	(-0.82, 0.582)
CV_Gap Duration	9	0.00152 (0.0015)	15	0.000986 (0.0015)	0.454	(-0.546, 1.15)
CV_Mean Peak Frequency	**9**	**0.0499 (0.0408)**	**15**	**0.0099 (0.0408)**	**1.37**	**(0.801, 1.77)**
CV_Song Duration	9	0.117 (0.139)	15	0.151 (0.139)	-0.199	(-0.829, 0.774)
Dom Frequency Range	12	2.21 (0.303)	15	2.16 (0.303)	0.172	(-0.608, 0.958)
Element Count	**12**	**15.4 (6.28)**	**15**	**43.6 (6.28)**	**-3.79**	**(-1.88, -1.65)**
***Element Diversity**	**12**	**1.53e-05 (8.21e-06)**	**15**	**3.44e-05 (8.21e-06)**	**-3.11**	**(-1.88, -1.53)**
***Element Duration**	**12**	**0.0667 (0.00862)**	**15**	**0.0615 (0.00862)**	**0.778**	**(0.0664, 1.28)**

Element Rate	12	10.3 (0.914)	15	10.7 (0.914)	-0.516	(-1.16, 0.27)
***Entropy**	**12**	**0.831 (0.0121)**	**15**	**0.792 (0.0121)**	**3.27**	**(1.54, 1.85)**
***Frequency IQR**	**12**	**0.916 (0.189)**	**15**	**1.04 (0.189)**	**-0.816**	**(-1.45, -0.0324)**
***Frequency Range**	**12**	**7.11 (2.49)**	**15**	**9.46 (2.49)**	**-1.2**	**(-1.77, -0.264)**
Gap Duration	12	0.0347 (0.00477)	15	0.0338 (0.00477)	0.21	(-0.583, 0.994)
***Mean Peak Frequency**	**12**	**3.37 (0.299)**	**15**	**3.64 (0.299)**	**-0.881**	**(-1.45, -0.125)**
Song Duration	**12**	**1.49 (0.543)**	**15**	**4.12 (0.543)**	**-3.45**	**(-1.85, -1.62)**
***Tail Length**	**15**	**80.2 (5.36)**	**15**	**90.5 (5.36)**	**-1.21**	**(-1.45, -0.563)**
Throat Brightness	13	22.2 (5.81)	14	20.5 (5.81)	0.3	(-0.438, 1)
Throat Chroma	13	0.482 (0.0443)	14	0.486 (0.0443)	-0.0779	(-0.825, 0.644)
Throat Hue	13	641 (33.9)	14	654 (33.9)	-0.458	(-1.09, 0.327)
Vent Brightness	15	27.7 (7.2)	14	23.5 (7.2)	0.651	(-0.0368, 1.19)
***Vent Chroma**	**15**	**0.454 (0.0413)**	**14**	**0.487 (0.0413)**	**-0.725**	**(-1.21, -0.0361)**
Vent Hue	15	628 (23.4)	14	645 (23.4)	-0.642	(-1.28, 0.0747)

Supplementary Methods and Background

Clarifying notes on female song

Popular field guide entries for barn swallows do not specify which sex sings, and it can be assumed the song description refers to males (Sibley, 2014; Stokes & Stokes, 2010). In contrast, the *Peterson Field Guide to Bird Sounds of Eastern North America* states "females also reported to sing" (Pieplow, 2017) and the entry in *Birds of the World* explicitly states that both sexes sing the "twitter-warble" song, i.e. typical male song (Brown & Brown, 2020). These latter interpretations may stem from Samuel's description of the barn swallow vocal repertoire, in which he states that the courtship song "is occasionally given by the female" (Samuel, 1971). He also describes a "whine call," which may be the same as the copulation call ("waeae-waeae") referred to in the *Birds of the World* (Brown & Brown, 2020), which may refer to what we call song in this study. In lieu of the original recordings or a high quality spectrogram, it is difficult to assess which female vocalizations are being described in previous studies. In Møller's *Sexual Selection and the Barn Swallow* (1994) there is no mention of female song or vocalizations at all. In fact the copulation call is ascribed to the male (p78). We believe this vocalization, which we transliterate as "dwidle dwight," is more accurately described as a song rather than a call (sensu Marler, 2004), since it is multisyllabic, structurally complex, and seems to have a function in mate attraction, pair bonding, and competition, based on observations (see supplementary video).

General acoustic analysis details

Prior to song analysis, all songs were standardized to a sample rate of 44.1 kHz and a bit depth of 16. We used Raven 1.5 (Bioacoustics Research Program, 2011) to select and extract measurements for each element in each song. We defined an element as a single, continuous trace on a spectrogram, separated from other elements by a visible break in time or an extreme shift in overall harmonic structure (occasionally, structurally distinct barn swallow elements were separated by very short breaks in time, so this additional criterion helped us distinguish among elements). For pure tone elements, we selected and compared the fundamental frequency of the element, whereas for harmonically rich elements, we selected the fundamental to the dominant frequencies. We did substantial visual checking to be sure that elements were separated and measured consistently across all songs. We used the following Hann window settings with a sample size of 512 (DFT), hop size of 256, and 50 percent overlap in Raven v1.5 for selecting and measuring elements (Bioacoustics Research Program, 2011).

Element-level acoustic parameters

For every element in all songs we extracted a large number of energy-based measurements calculated from the cumulative amplitude of the spectrum or envelope using Raven Bioacoustic Analysis Software v1.5 and the warbleR package in R (Araya-Salas & Smith-Vidaurre, 2017; Bioacoustics Research Program, 2011). We removed any highly correlated variables with a Pearson correlation coefficient greater than 0.95. This resulted in the following acoustic variables measured for each element using each bioacoustics program: Raven measurements – peak frequency, frequency 5%, delta frequency, interquartile range bandwidth, time 5%, interquartile range duration, duration 90%, aggregate entropy, average entropy, average power, energy; warbleR measurements – standard deviation of frequency, mean dominant frequency, mean peak frequency, minimum dominant frequency, maximum dominant frequency, 1st quartile frequency, interquartile frequency range, dominant frequency range, dominant frequency slope, 1st quartile time, spectral entropy, time entropy, entropy, spectral flatness, modulation index, skewness, and kurtosis. See Raven and warbleR manuals for definitions of each parameter (Araya-Salas & Smith-Vidaurre, 2017; Bioacoustics Research Program, 2011).

Song-level acoustic parameters – Using the warbleR package in R, we extracted the following acoustic parameters for each song: mean frequency, mean peak frequency, median frequency, frequency range, dominant frequency range, 1st quartile frequency, 3rd quartile frequency, interquartile frequency range, median time, 1st quartile time, 3rd quartile time, entropy, average element duration per song, song duration, number of elements, element rate, and gap duration. See the warbleR vignette for descriptions of each parameter (Araya-Salas & Smith-Vidaurre, 2017).

Element diversity calculations – Using the element-level acoustic parameters, we estimated element diversity for each song as the area encompassing the elements for that song in a multi-dimensional feature space. To do this, we first created an element-level acoustic space by inputting the element-level acoustic parameters for every element for all songs into an unsupervised random forest. We used the 'randomForest' function in the R package {randomForest} with the following specifications: 10000 trees, minimal node size of 1, Gini impurity index as split rule, 5 randomly sampled variables at each split and out-of-bag proximity (Liaw & Wiener, 2002). This produced a proximity matrix, which we transformed into a set of five vectors using classic multidimensional scaling, using the 'cmdscale' function in the {stats} R package. We then calculated and extracted minimum convex polygons to calculate the area surrounding all elements for each song within this multidimensional space using the function 'mcp' in the {adehabitatHR} R package (Calenge, 2006). Therefore, each area is an estimate of the variability, or diversity, of elements for a given song within an acoustic feature space of all barn swallow element measurements.

Supplemental References:

Araya-Salas, M., & Smith-Vidaurre, G. (2017). warbleR: anrpackage to streamline analysis of animal acoustic signals. In Methods in Ecology and Evolution (Vol. 8, Issue 2, pp. 184–191). https://doi.org/10.1111/2041-210x.12624

Bioacoustics Research Program. (2011). Raven Pro: Interactive Sound Analysis Software (Version 1.5) [Computer software]. http://www.birds.cornell.edu/raven

Brown, M. B., & Brown, C. R. (2020). Barn Swallow - Hirundo rustica. Birds of the World. https://doi.org/10.2173/bow.barswa.01

Calenge, C. (2006). The package “adehabitat” for the R software: A tool for the analysis of space and habitat use by animals. In Ecological Modelling (Vol. 197, Issues 3-4, pp. 516–519). https://doi.org/10.1016/j.ecolmodel.2006.03.017

Liaw, A. L., & Wiener, M. (2002). Classification and regression by randomForest. R News, 2, 18–22.

Marler, P. (2004). Bird calls: a cornucopia for communication. In Nature’s Music (pp. 132–177). https://doi.org/10.1016/b978-012473070-0/50008-6

Møller, A. P. (1994). Sexual Selection and the Barn Swallow. Oxford University Press.

Pieplow, N. (2017). Peterson Field Guide to Bird Sounds of Eastern North America. Houghton Mifflin Harcourt.

Samuel, D. E. (1971). Vocal Repertoires of Sympatric Barn and Cliff Swallows. In The Auk (Vol. 88, Issue 4, pp. 839–855). https://doi.org/10.2307/4083842

Sibley, D. A. (2014). The Sibley Guide to Birds. Knopf.

Stokes, D., & Stokes, L. (2010). The Stokes Field Guide to the Birds of North America. Little, Brown.