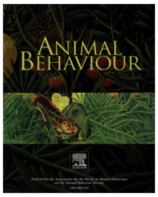
Animal Behaviour 168 (2020) 69–82

Contents lists available at ScienceDirect

Animal Behaviour

journal homepage: www.elsevier.com/locate/anbehav



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

Matthew R. Wilkins [a], [b], [*], Karan J. Odom [c], [d], Lauryn Benedict [e], Rebecca J. Safran [a]

[a] Department of Ecology and Evolutionary Biology, University of Colorado at Boulder, CO, U.S.A.
[b] Collaborative for STEM Education and Outreach, Vanderbilt University, Nashville, TN, U.S.A.
[c] Department of Neurobiology and Behavior, Cornell University, Ithaca, NY, U.S.A.
[d] Cornell Lab of Ornithology, Cornell University, Ithaca, NY, U.S.A.
[e] School of Biological Sciences, University of Northern Colorado, Greeley, CO, U.S.A.

ARTICLE INFO

Article history:
Received 25 March 2020
Initial acceptance 14 May 2020
Final acceptance 22 June 2020

MS. number: A20-00197R

Keywords:
animal communication
barn swallow
female song
Hirundo rustica
multimodal signalling
sexual dimorphism
sexual selection

Understanding the patterns and processes related to sexual dimorphism and sex differences 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 sex differences via elaboration of male traits, but selection on females also has the potential to drive divergence. Here, we describe female song in barn swallows, *Hirundo rustica erythrogaster*, for the first time, report rates of female song production and couple song data with plumage data to explore the relative degree to which sex differences in phenotypic traits are consistent with contemporary selection on males versus females. During previous intensive study of male song over 2 years, we opportunistically recorded songs for 15 females, with matched phenotypic and fitness data. We randomly selected 15 high-quality samples from our larger male data set to test whether sex differences in song and plumage are 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 were used in multiple contexts, restricted primarily to the beginning of the breeding season; song traits showed greater sexual differentiation 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 differentiation. To achieve a better understanding of sex differences and dimorphism, we require comprehensive studies that measure the same traits in males and females and their fitness consequences.

Ecologists and evolutionary biologists have long sought to understand the processes driving dimorphism and other sex-based phenotypic differences (Andersson, 1994; Badyaev & Hill, 2003; Burns, 1998; Darwin, 1801, p. 396; Darwin, 1859, p. 94; Hedrick & Temeles, 1989; Lande, 1980; Ng et al., 2019). However, owing to historical biases, studies of the drivers of differentiation in sexual signalling traits have traditionally focused on male signals, and most approaches assume that sexual selection has promoted sex differences via elaboration of male traits (Badyaev & Hill, 2003; Freed, 2000; Langmore, 1998; Riebel, 2016; Riebel, Hall, & Langmore, 2005; Riebel, Odom, Langmore, & Hall, 2019; Rosvall, 2011). Yet, meta-analyses of the strength of sexual selection on male traits report moderate effect sizes (Jennions, Kahn, Kelly, & Kokko, 2012), with evidence that males are often under variable selection pressures within and across breeding seasons (Chaine & Lyon, 2008; Kingsolver, Diamond, Siepielski, & Carlson, 2012; Robinson, Pilkington, Clutton-Brock, Pemberton, & Kruuk, 2008; Steele, Siepielski, & McPeek, 2011), have trait values near optima (Evans, 1998; Rodríguez, Ramaswamy, & Cocroft, 2006), or that the magnitude of trait differences may not reflect the strength of current selection on males (Miller, McDonald, & Moore, 2016). Indeed, sex differences 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, 2012; Dunn, Armenta, & Whittingham, 2015; Price, 2015; Shultz & Burns,

---

[*] Correspondence: M. R. Wilkins, Collaborative for STEM Education and Outreach, Vanderbilt University, PMB 0367, 230 Appleton Pl., Nashville, TN, 37212, U.S.A.
E-mail address: m.r.wilkins06@gmail.com (M. R. Wilkins).

https://doi.org/10.1016/j.anbehav.2020.07.018
0003-3472/

2017; Wiens, 2001). To better understand how divergent selection operating between the sexes drives sex differences, we require holistic approaches that fully describe the traits of both males and females (Hare & Simmons, 2019; Riebel et al., 2019). Ideally, studies should include multiple signalling traits that mediate inter- and intrasexual interactions (Bro-Jørgensen, 2010; Hebets et al., 2016; Hebets & Papaj, 2004; Partan & Marler, 2005), as well as their fitness consequences.

Birdsong and plumage colour offer excellent examples for how female, as well as male, signal evolution can lead to sex differences. Females sang in the ancestor of modern songbirds and still sing in many tropical and temperate oscines (Garamszegi, Pavlova, Eens, & Møller, 2006; Odom, Hall, Riebel, Omland, & Langmore, 2014; Price, Lanyon, & Omland, 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 colour in both males and females (Dale, Dey, Delhey, Kempenaers, & Valcu, 2015; Hofmann, Cronin, & Omland, 2008; Price & Eaton, 2014), with females experiencing more evolutionary plumage change than males in some lineages (Price & Eaton, 2014). Even though losses of female song and elaborate plumage appear to contribute considerably to sex differences in signalling 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, Roper, & Harmer, 2016; Kleindorfer, Evans, & Mahr, 2016).

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, Hebets, & Podos, 2010; Møller, 1988; Rodríguez et al., 2006; Ryan et al., 2019; Searcy & Andersson, 1986). In contrast, relatively few studies have looked for or uncovered relationships between female visual signals and fitness proxies (but see Amundsen, Forsgren, & Hansen, 1997; Bulluck et al., 2017; Cain & Ketterson, 2011; Jawor, Gray, Beall, & Breitwisch, 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, Cockburn, & Langmore, 2015; Kleindorfer et al., 2016; Krieg & Getty, 2016). Although inclusive, systems-based approaches can best untangle the relative contributions of selection on males and females to signal evolution across modalities (Hebets et al., 2016; Riebel et al., 2019), researchers have seldom included both visual and acoustic phenotypes from both sexes within the same study (Hebets et al., 2016; Riebel et al., 2019; but see Webb et al., 2016). Moreover, to our knowledge, no study has directly tested whether the pattern of sex differences in vocal and visual sexual signalling 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 data sets 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 erythrogaster*, offer a valuable system in which to investigate the selection pressures governing the evolution of sexual signalling traits and associated sex differences. 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 sex differences are more strongly associated with contemporary selection (i.e. trait correlations with reproductive success) in males or females. Because the differences between several measured male and female traits were not categorical, we conservatively discuss them as sex differences rather than dimorphism although other studies of this system may choose to call them dimorphic.

Much previous research is predicated on the idea that sex differences and dimorphism in communicative traits result from selection for male trait exaggeration (Badyaev & Hill, 2003; Freed, 2000; Langmore, 1998; Riebel, 2016; Riebel et al., 2005, 2019; Rosvall, 2011). 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 differentiation in a variety of visual and acoustic traits (similar to Badyaev & Martin, 2000). If sex differences are driven/maintained by directional selection on males (for showier or more elaborate traits), then we expected that trait correlations with male reproductive success would predict degree of sexual differentiation. However, if differences are driven/maintained by directional selection on females (for more cryptic or energy-efficient communication), then we expected that trait associations with female reproductive success would predict levels of sexual differentiation. While the combined effect of differential selection on the sexes is ultimately responsible for the overall pattern of sex differences (Hedrick & Temeles, 1989; Price, 1984), we are chiefly concerned with testing the common assumption that it is sexual selection on males, rather than females, that drives this pattern in signalling traits. Alternatively, if trait sex differences arose in the past, but are currently under stabilizing selection, we would expect to find that the magnitude of trait sex differences have no contemporary associations with reproductive performance.

METHODS

Background

As the subject of hundreds of studies on sexual signal evolution since the late 1980s, the barn swallow offers an excellent system for testing questions related to sex differences. Barn swallows are a weakly dimorphic species, in which males have darker ventral plumage and longer tail feathers on average. They comprise six Holarctic subspecies that have rapidly diversified from a common ancestor as recently as 7700 years ago (Smith et al., 2018). In North American barn swallows, males with darker ventral feathers have higher reproductive success, while tail feather length is not a preferred trait in males (Eikenaar, Whitham, Komdeur, van der Velde, & Moore, 2011; Safran et al., 2016; Safran & McGraw, 2004), or females (Safran & McGraw, 2004). There is limited evidence of selection for darker plumage in females in a New York population of *H. r. erythrogaster* (Safran & McGraw, 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, Navarro, de Lope, & Møller, 2008; Galeotti, Saino, Sacchi, & Møller, 1997; Galeotti, Saino, Perani, Sacchi, & Møller, 2001; Garamszegi, Heylen, Møller, Eens, & de Lope, 2005; Garamszegi, Hegyi et al., 2006; Saino et al., 2003; Wilkins, Shizuka, Joseph, Hubbard, & Safran, 2015). Despite intensive study of this species, female signals have received relatively less 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, 2020), 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 in Boulder County, Colorado, U.S.A., which was set up in 2008 by M.R.W. by contacting local birders and equestrian clubs, surveying the county for culverts and structures with old nests, and, after arrival, looking for signs of flying or perched swallows nearby. The five sites included in this study covered roughly 65 km^2 and included 8–64 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 colour bands, and application of Sharpie marker ink combinations to left and right rectrices). These individual markings were used to identify individuals during sound recording. At the time of capture, tail streamer length was measured and a small sample (<90 μl) of blood was taken via brachial venipuncture and stored in 2% lysis buffer for later parentage analysis. Additionally, a set of approximately five contour feathers was sampled along a ventral transect (throat, breast, belly, vent) and attached to index cards with tape for colour analysis using a spectrophotometer in the laboratory (Hubbard et al., 2017; Wilkins et al., 2015). Nests were closely monitored for egg laying activity, with checks conducted at least every 4 days to determine clutch initiation date and date of hatching. On approximately day 12 posthatching, all nestlings in a nest were banded and blood samples were taken for parentage analysis.

Ethical Note

We made every possible attempt to minimize handling time and other sources of stress on swallows in our study population. Capture typically involved closing off all but one exit from a barn or other structure, in front of which a fine mesh nylon mist net was held or attached in place with bungee cords. Nets were constantly monitored and birds were usually removed within 1 min of landing in the net and placed into custom-sewn cloth bags (closed with a ribbon) to calm them until they could be transported to a temporary banding station a short distance away (usually <30 m), sampled and released. Birds held temporarily in bags were kept in a quiet, cool location until they were removed for processing. Typically, birds remained quiet and fairly still both in the net and while in the bag. Obvious signs of stress during handling were attended to constantly. In the rare instances where a bird showed signs of distress during netting, banding, measurement or blood sampling, it was immediately released. 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

Consistent with a previous study on complex male signalling in this population (Wilkins et al., 2015), we measured maximum tail streamer length because it is more normally distributed than mean tail streamer length (see Appendix Table A1 for trait definitions). Additionally, we measured a total of 12 traits to characterize colour variation: three axes of colour (average brightness, hue, red chroma) along a four-patch ventral transect (throat, breast, belly, vent). We opted to use these hue–saturation–brightness (HSB) measures to make interpretation and comparison to other studies in this system as easy as possible. We are also confident about the biological relevance of these trait measures, since they strongly correlate with eumelanin and phaeomelanin concentrations in feathers (McGraw, Safran, & Wakamatsu, 2005); ventral feathers do not show an ultraviolet (UV) reflectance peak (Safran & McGraw, 2004); and colour manipulations (measured using HSB) to mimic natural melanization profiles have shown predictable impacts on physiology and reproductive performance (Safran, Neuman, McGraw, & Lovette, 2005; Safran, Adelman, McGraw, & Hau, 2008; Safran et al., 2016). However, as a check, we calculated colour measures from raw spectra in the tetrahedral colour space (TCS) that incorporates a model of avian visual sensitivity (Endler & Mielke, 2005; Stoddard & Prum, 2008). See Supplementary Methods for details on how we calculated both HSB and TCS colour traits. As shown in the Appendix, Fig. A2, average brightness is exactly equivalent to the TCS measure 'brilliance' described in Stoddard and Prum (2008). Our measures of chroma showed a strong positive (0.84) Spearman's correlation with the equivalent TCS measure of saturation (rA, achieved chroma). Hue showed a moderate (−0.35) correlation with brilliance, although direct comparison with TCS measures of hue is challenging because TCS θ and Φ jointly describe stimulation of avian cones and a single correlation between hue and one of these measures in not meaningful. Overall, given biochemical and experimental evidence and a lack of significant UV reflectance in barn swallow ventral coloration (Hubbard et al., 2017; McGraw et al., 2005), we are very confident that our HSB measures of colour capture biologically meaningful variation in this system.

Fitness Measures

We used female seasonal (i.e. annual) fledging success and male within-pair genetic paternity as fitness proxies for females and males, 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 fledged genetic offspring within the social nest across all broods, determined through genetic paternity exclusions. Paternity assignments (and total extrapair paternity) were not feasible in our study area, due to the presence of known, unmonitored breeding sites within easy flying distance from monitored sites. To get a representative sample of males' extrapair young (EPY) would also require genotyping thousands of individuals in any given year. Briefly, genotypes were derived from fluorescently labelled PCR products of six microsatellite loci. Paternity was conducted using CERVUS v.3.03 (Kalinowski, Taper, & Marshall, 2007) and an offspring was considered extrapair when the offspring–mother–father trio confidence did not reach the 95% level. Detailed paternity exclusion methods are provided in the Supplementary Methods. 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 opportunistically as part of a study of male song (Wilkins et al., 2015) conducted in 2011 and 2012. Although males were the target of the prior study, this rarer, unpredictable female vocalization was of growing interest to M.R.W. When a female vocalization was heard, M.R.W. immediately redirected the microphone at the source of the sound, capturing as many vocalizations as possible. Given the complex acoustic environment in which barn swallows live, active, annotated recording is necessary for confident singer identity. Thus, our samples by no means capture comprehensive singing outputs for individuals; however, our estimates of relative singing outputs should reflect realistic activity patterns. Songs were recorded in 16-bit WAV format, with 48 kHz sampling rate using a Marantz PMD 660 paired with a Sennheiser MKH 20 and Telinga parabola (in 2011), or a Marantz PMD 661 paired with a Sennheiser ME62/k6 microphone and Telinga parabola (in 2012). Total recording time was approximately 57 h (2–31 May) in 2011 and 48 h (1 May – 21 August) in 2012, all between 0500 and 1200 hours (Wilkins et al., 2015).

Acoustic Trait Measures

Using Raven v.1.5 (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 males (chosen randomly from among those with high-quality song recordings) to create two acoustic data sets: (1) an element level data set of acoustic parameters measured for every element in all songs and (2) a song level data set of acoustic parameters averaged for all elements in a song or calculated for the whole song. The element level data set was used to estimate an element diversity score (Keen et al., 2020), which estimates the variability, or diversity, of elements for a given song within an acoustic feature space of all barn swallow element measurements (see full explanation in Supplementary Methods). This value was then appended to the song level data set, 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 signalling, based on previous studies of male song (Galeotti et al., 1997; Garamszegi, Hegyi, et al., 2006; Wilkins et al., 2015). The 11 selected acoustic variables 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 Appendix Table A1 for definitions). We then calculated individual means and coefficients of variation for each acoustic variable. This resulted in 22 variables describing differences in the magnitude and variability of different aspects of song frequency, timing and elemental diversity for both sexes. The number of song traits was further reduced to avoid redundancy in analysis of sex differences 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. Our subsample of males was temporally equivalent to that for females, as the sexes did not differ in recording date relative to first clutch initiation (mean ± SD: females: -6.6 ± 10.7; males: -2.6 ± 21.1; unpaired t test: $t_{20.8} = -0.65$, $P = 0.52$). 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 using sample sizes of 9–15 individuals. Although our sample sizes were 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 v.3.6.0 (R Core Team, 2019). To measure the magnitude of the sex difference 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 difference values indicate male-biased differences and positive values indicate female-biased differences. For example, in barn swallows, tail feathers show male-biased differentiation (as males have longer tails than females on average), resulting in a negative value on our sex difference scale. To determine which traits were robustly distinct, we created 10 000 bootstraps using the 'bootstraps' function in the 'rsample' package v.0.0.5 (Kuhn, Chow, & Wickham, 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 variables with Spearman's rank correlations $>|0.7|$ and retained the most biologically intuitive variables for analysis. 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 nondimorphic traits, we were left with 10 visual and acoustic traits for exploring the connection between sex differences and contemporary selection pressures.

Testing traitwise associations between reproductive success and sexual differentiation

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 our surrogate measures of selection and observed trait sex differences, while accounting for nonindependence of traits, we used the 'crunch' function in the 'caper' package v.1.0.1 (Orme et al., 2018), which implements the CAIC (comparative analysis by independent contrasts) algorithm (Purvis & Rambaut, 1995). Following a similar approach to Garamszegi, Pavlova, 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 $1 - |(\text{Spearman's correlation})|$, and constructed the tree using a single-linkage clustering method with the base 'hclust' function.

Thus, we made a linear model with a trait's magnitude of sex difference as a response and a trait's correlation with reproductive success as a predictor, controlling for correlation structure among traits. Given our metric of sexual differentiations, we would expect a significant positive association between a trait's correlation with reproductive success and the magnitude of sexual differentiation if female fitness is maintaining sex-based trait differences. That is, traits that confer fitness benefits to females should have a high degree of differentiation for our metric, and traits that reduce fecundity in females should have low differentiation values. We would expect the opposite (a negative slope) if male fitness is driving/maintaining sexual differentiation. That is, traits associated with higher paternity would be more exaggerated in males (having more negative differentiation values on our metric), and traits that reduce paternity would be less exaggerated in males (higher differentiation values), resulting in a negative slope.

Data availability

Male and female song clips, HSB and TCS colour, and other phenotypic and fitness data are available in our GitHub repository at https://github.com/drwilkins/femaleSongInBARS. Spectrograms in Fig. 1 were generated with R code, available at https://github.com/drwilkins/rspect.

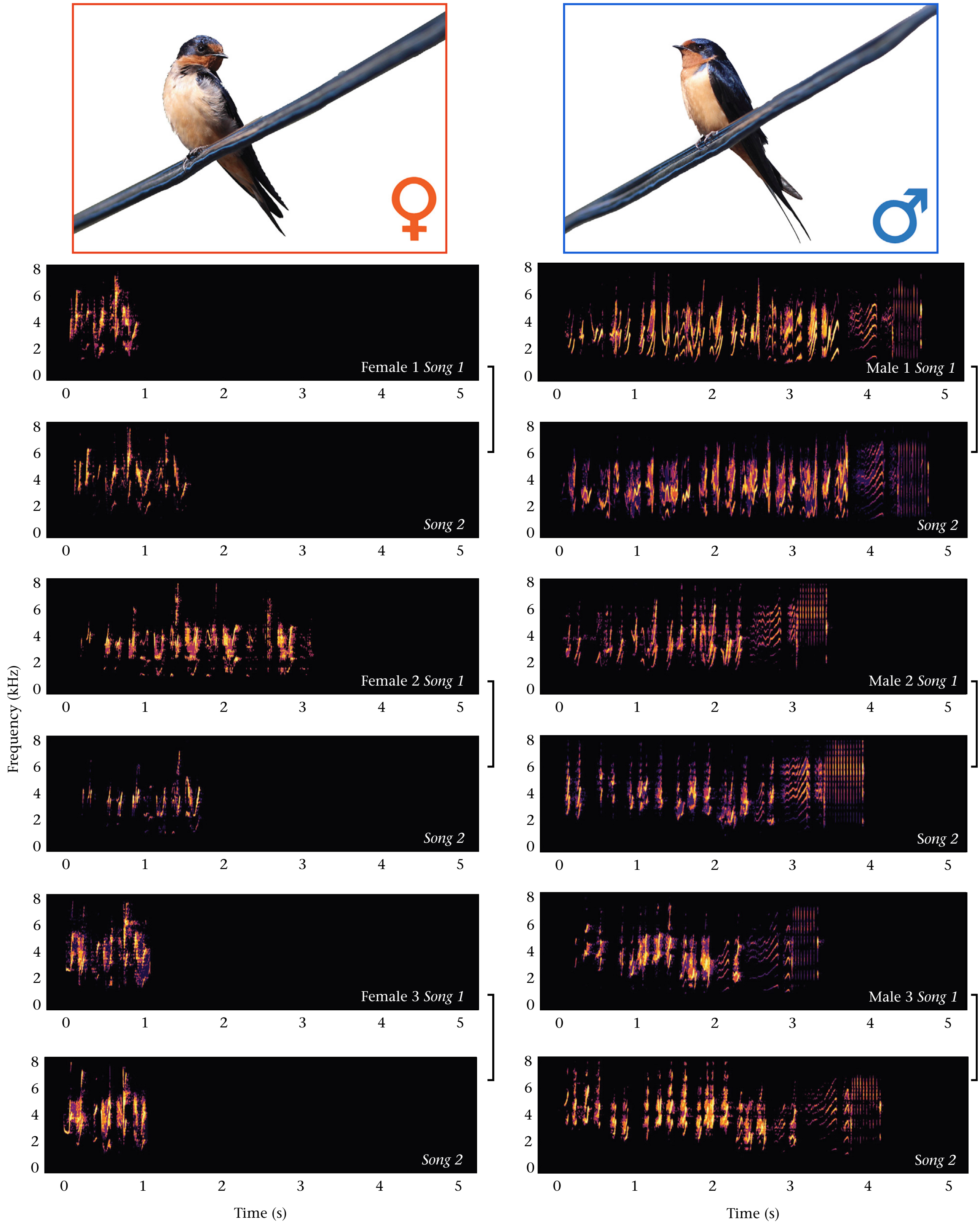


Figure 1. Spectrograms of female and male song renditions in barn swallows. Two song renditions are shown for three females and three males to demonstrate variability. Photos by M.R.W.

RESULTS

Characterizing Female Song

Recordings and observations revealed that females used a distinctive vocalization in spontaneous solo renditions, and in response to hearing other females, in a manner synonymous to male countersinging. Females were also observed using the vocalization to interrupt the songs of their mates (see video at https://vimeo.com/424642268). This vocalization is relatively complex, including many syllables with similar frequency modulation and pace to male song (Fig. 1). Collectively, this evidence confirms that female barn swallows do produce facultative song, rather than just simple calls (Langmore, 1998).

Female song bouts were often produced within the nest, but also on perches near the nest or outside the barn/structure, and were usually short and infrequent. As such, only 78 songs were recorded from 18 identifiable females over the course of 2 years, compared to 753 (865% more) songs from 40 males, given the same recording effort (~105 h). This amounts to a recording efficiency of about 7.17 clear, identifiable songs/h for males and 0.74 songs/h 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 1 May and 21 August, while female songs were only observed and recorded between 10 May and 29 May (pooling years) (Fig. 2a). Restricting to only the active singing period for females during May of both years, a cumulative 91 h of recording documented 8.9 individually identifiable songs/h for males and 0.86 songs/h for females. We could not calculate exact song rates per individual due to the complexities of recordings with shifting focal singers in dynamic colonial environments. To explore effects of breeding phenology on singing output, we also calculated relative recording dates by subtracting the ordinal date of first clutch initiation (i.e. breeding onset) from the ordinal recording date. For males, these relative recording dates ranged from 59 days before to 109 days after the first egg was laid by a male's mate, with an average date (weighted by number of songs recorded each day) of 5.43 days after breeding onset; Fig. 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 resulting from low recording quality, the total number of songs analysed was 48 for females (mean ± SD = 4.0 ± 3.02, range 1–11) and 66 for males (4.4 ± 0.99, range 2–5).

Visual and Acoustic Differentiation in Males and Females

Our bootstrap analysis of 35 acoustic and visual traits showed that 16 were robustly distinct in the two sexes – i.e. had bootstrap 95% confidence intervals that did not overlap zero (Fig. 3). These included a mixture of song and visual traits, with song traits being most distinctive (13/22 = 59% for acoustic traits, compared to 3/13 = 23% for visual traits). Traits that had higher values in females (showing female-biased differences) included entropy, CV_mean peak frequency, CV_dominant frequency range, CV_frequency IQR, CV_element diversity, belly brightness and element duration. Male-biased dimorphic traits included vent chroma, frequency IQR, mean peak frequency, CV_entropy, frequency range, tail length, element diversity, song duration and element count. Most notably, female songs were brief, they included few elements and included elements with high entropy (i.e. noise) and they had low element diversity across songs, compared to males (Fig. 3). Consistent with previous results for visual traits (Safran & McGraw, 2004), females had shorter tail streamers and lighter ventral plumage (higher brightness scores for belly and lower chroma scores for the vent).

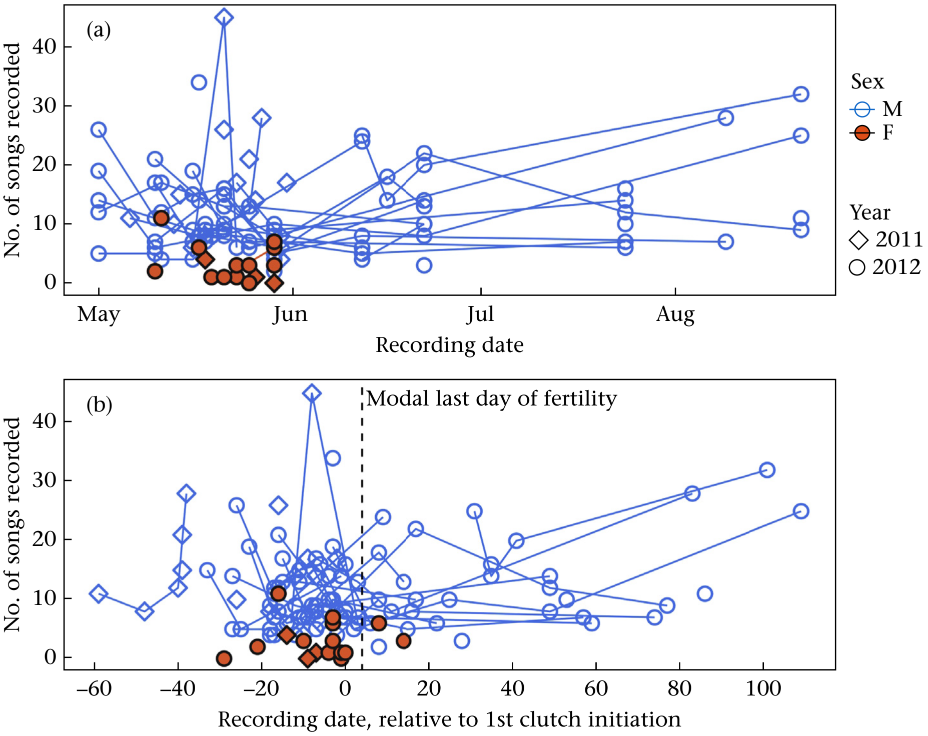


Figure 2. Seasonal variation in song production, as demonstrated by opportunistic recording of spontaneous male and female songs. (a) Number of songs recorded for both sexes as a function of recording date. (b) Number of songs recorded for both sexes as a function of breeding onset (ordinal recording date minus ordinal date of breeding onset; first clutch initiation); negative values on the *X* axis represent recordings taken before breeding onset. The dashed vertical line marks day 4 after clutch initiation–i.e. the last day of female fertility before laying of the ultimate egg of a modal five-egg clutch. Lines connect points for individuals with recordings on multiple days.

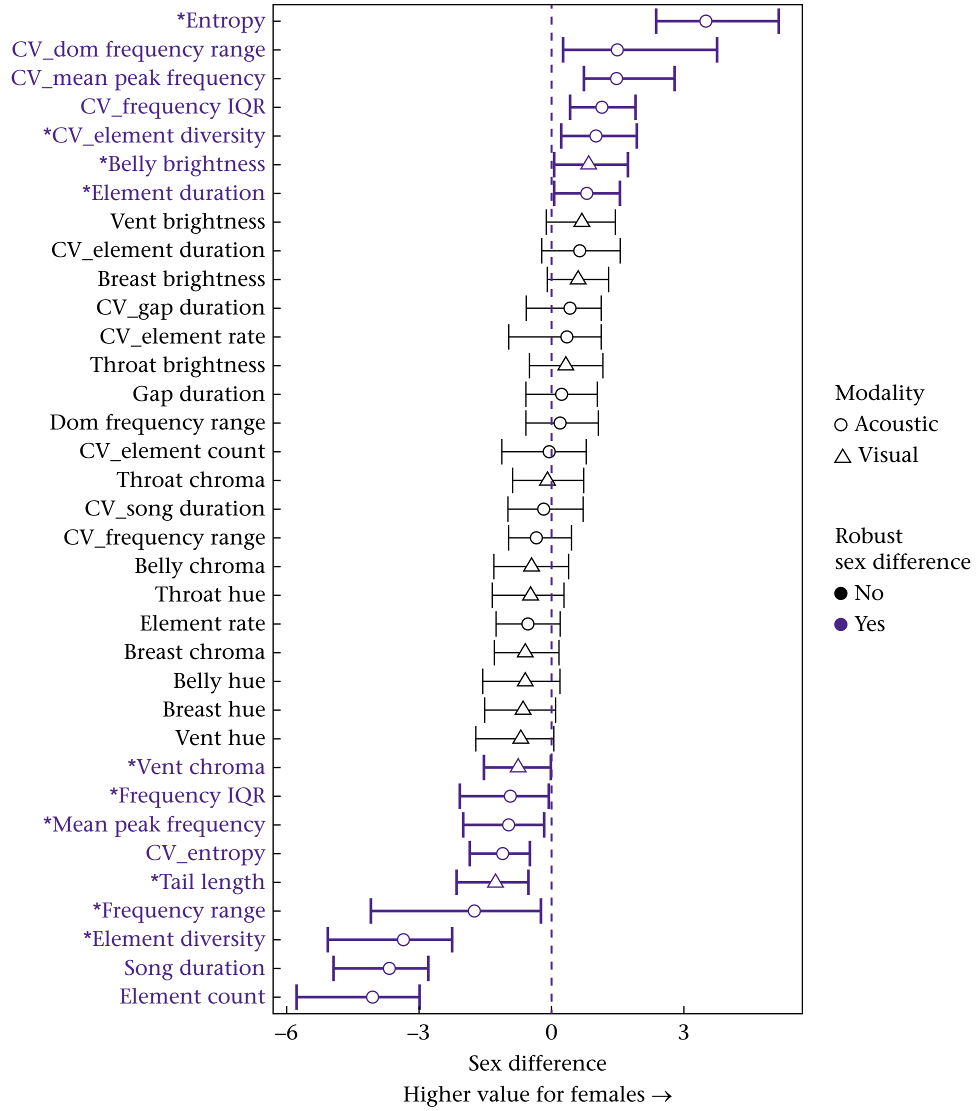


Figure 3. Sex differences in 35 potential sexual signalling traits. Empirical point estimates for Cohen's *d* are shown with 95% bootstrapped confidence intervals; intervals not overlapping zero are shown in bold. Traits are means, except where preceded by CV (coefficient of variation). *Denotes the chosen set of robustly sexually distinct traits used for further analysis, after removing highly redundant traits with Spearman's correlations $\geq |0.7|$.

Trait means, sample sizes and sex difference confidence intervals are reported in the Appendix, Table A2.

Connecting Contemporary Selection to Phenotypic Sex Differences

Our final objective was to test whether sexual differentiation is more strongly associated with contemporary selection in males or females (here, using surrogate measures of lifetime fitness: within-pair reproductive performance for a single breeding season). After controlling for intercorrelations between the 10 nonredundant, sexually distinct traits using independent contrasts, we found a significant positive relationship between trait sex differences and our measure of current selection pressures (i.e. a correlation with either seasonal fledging success or genetic paternity) in females (slope = 6.04, $t_8 = 2.92$, $P = 0.019$), but not in males (slope = 1.06, $t_8 = 0.547$, $P = 0.599$) (Fig. 4). That is, traits that had a negative association with female reproductive success (see Appendix, Fig. A1) showed male-biased sex differences (i.e. greater trait expression in males), while traits that positively correlated with female reproductive success showed female-biased sex differences. The relationship remained significant for females when including all 16 robustly distinct traits (slope = 4.93, $t_{14} = 3.53$, $P = 0.003$), as well as for the subset of distinct acoustic traits (slope = 5.28, $t_{11} = 3.70$, $P = 0.003$). In contrast, the relationship between male traits and within-pair genetic paternity did not predict levels of sex differences for either of these subsets (all $P \geq 0.260$).

DISCUSSION

Female Song in Barn Swallows

Despite intensive study of this species for decades, female song in barn swallows has been reported as absent by some authors (Sibley, 2014; Stokes & Stokes, 2010), the same as male song by others (Brown & Brown, 2020), or left ambiguous (Pieplow, 2017) (see Supplementary Methods for notes on previous descriptions). Our data and videos confirm that female barn swallows sing a quantitatively distinct song from males. We report that while female song is approximately 10 times less frequent than male song and relegated largely to the first month of the breeding season, it

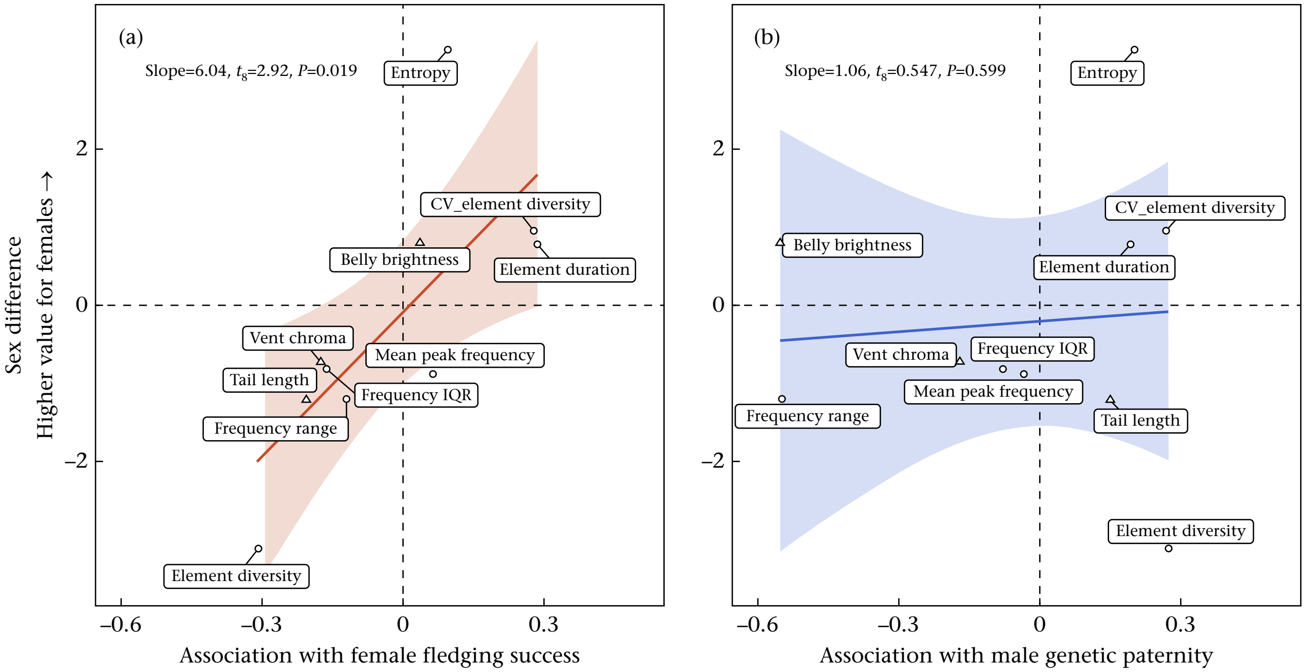


Figure 4. Relationship between trait sexual differentiation and trait association with reproductive output for (a) females and (b) males. Sexual differentiation was measured as the difference of scaled female and male trait values in pooled standard deviation units (Cohen's d). Reproductive output was measured as seasonal fledging success for females and fledged within-pair seasonal genetic paternity for males.

has an overall similar structure to male song and seems to be used similarly in countersinging, as well as mate interruption contexts. The use of female song in countersinging (see video at https://vimeo.com/424642268) just prior to clutch initiation is consistent with singing patterns and use of song for territory/nest defence or intrasexual competition in other northern temperate breeding species (Cain et al., 2015; Krieg & Getty, 2016; Levin, 1996; Magoolagan, Mawby, Whitehead, & Sharp, 2019; 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, 2016; Rosvall, 2011). Furthermore, 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 & Bitz, 2004; Tobias & Seddon, 2009).

Although female and male barn swallow songs have a generally similar form, female songs lack the terminal trill that is universal for male song among different barn swallow subspecies (Wilkins et al., 2018). Our results further indicate that female songs are shorter and noisier, with fewer elements, reduced frequency ranges and lower element diversity (Figs 3 and 4, Fig. A1). Thus, barn swallow song is categorically dimorphic in some aspects and shows variable sex differences in others. Overall, in this species, song traits are more sexually distinct (59% of measured traits) than plumage traits (23% of measured traits) and the direction of sex differences is predicted by both the trait correlations with female fledging success and qualitative expectations from selection for efficient, cryptic signalling in females. For example, element diversity (our measure of average syllable complexity, which was retained after removing the redundant measure 'song duration') was much lower in females and showed a strong negative relationship with fledging success. In contrast, element duration (the length of individual song elements) was significantly longer in females and showed the strongest positive correlation with reproductive success of any trait. Together, these results are consistent with selection for shorter songs with longer individual elements. Shorter songs may be favoured to avoid attracting nest predators, an effect previously shown for song rate (Kleindorfer et al., 2016). Shorter songs also offer the cognitive benefit of simpler comparison due to easier discriminability of proportional differences (Akre, Farris, Lea, Page, & Ryan, 2011; Akre & Johnsen, 2014).

Our finding that mean element diversity was male-biased and negatively correlated with female fledging success, but CV_element diversity (i.e. song versatility: variability of syllable complexity across songs) was female-biased and positively correlated with fledging success was unexpected. One possible explanation for this may stem from variable functions (and audiences) for female song. That is, signals in competitive contexts tend to be shorter and more repetitive, which would select for low element diversity if this is the primary function of song (Collins, de Kort, Pérez-Tris, & Tellería, 2009; Galeotti et al., 1997). On the other hand, song versatility and the ability to adaptively change element diversity across competitive and mating-related signalling contexts could be especially important for females if they are constrained to shorter signal rate and duration, and there is a premium on efficient information transfer about motivation or quality within a shorter time window than males. This is, of course, speculative and close study of the signalling context of song production is necessary to better understand these patterns of female song variation.

Taken together, our results provide a mechanism supporting the evolution of female song in many north temperate bird species from complex songs comparable to male songs towards highly condensed, context-dependent songs, or even towards the loss of songs altogether (Odom et al., 2014). This is significant because, while the vast literature on male birdsong (and acoustic signal evolution more broadly) has informed many aspects of sexual selection theory (Gil & Gahr, 2002; Nowicki, Searcy, & Peters, 2002), and male acoustic divergence has been shown to play a key role in premating isolation and speciation across diverse taxa (Alcaide, Scordato, Price, & Irwin, 2014; Blankers, Oh, & Shaw, 2019; Hasiniaina et al., 2020; Irwin, Bensch, & Price, 2001; Lee, Shaner, Lin, & Lin, 2016; Sosa-López, Martínez Gómez, & Mennill, 2016; Wilkins, Seddon, & Safran, 2013, 2018), we know relatively little about female vocal signalling or its implications for broader evolutionary processes, even in classic study systems like barn swallows.

Overall Patterns of Sexual Differentiation in Signals

In contrast to a common assumption in the sexual selection literature, we found that overall patterns of sex differences are better explained by (surrogate measures of) selection on females, rather than selection for more elaborate males. That is, the overall magnitudes of sex-based differences in song form, tail streamer length and colour were predicted by how those traits correlated with female, but not male, reproductive performance. This does not mean that selection on males (i.e. through female choice) is not important or relevant in explaining sex differences. For example, as shown in Fig. 4b and Fig. A1, belly brightness – a known sexual signal in males within this population (Safran et al., 2016; Wilkins et al., 2015) – showed a strong negative association with genetic paternity (selecting for darker males with lower brightness). In turn, belly brightness in females showed a slight positive association with fledging success, and sex difference was low (0.8 SD), but significantly biased towards lighter females. That is, higher reproductive performance among darker males, combined with negligibly higher performance in lighter females correlates with an overall pattern of slightly darker males. In contrast, a strong positive correlation between element diversity and reproductive success in males and a strong negative correlation for females (i.e. divergent selection in the sexes) corresponds with a highly male-biased sex difference (−3.11 SD) towards more complex syllable composition of songs. While an additional analysis testing the combined impact of traitwise reproductive performance on traitwise sex differences for both sexes is desirable, it would likely be similar to the result shown for females (Fig. 4a) and is not possible using our current comparative analysis by independent contrasts approach, as we are unaware of a way to control for known differences in trait correlation structure between the sexes in a single model.

Broader Implications

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, our 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 that rely on the degree of plumage differences as a surrogate for the strength of sexual selection. It is widely recognized that there are limitations to using sex differences, including dimorphism, as a surrogate for sexual selection (Bell & Zamudio, 2012; Huang & Rabosky, 2014; Kraaijeveld, Kraaijeveld-Smit, & Maan, 2011; Price, 2015). 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. Many studies assume that sex differences largely result from sexual selection for exaggerated male traits in conjunction with ecological selection for crypsis on females. This may be a valid assumption in some cases, such as in sexual dichromatism in damselflies (Svensson & Waller, 2013). However, our results highlight the need (especially in taxa with more complex communication and/or cognition) to consider how female competition and mutual mate choice affect the evolution of sex differences within the context of constraining ecological selection on females from increased costs of predation, migration and/or coloniality (Jawor et al., 2004; Price, 2015; Tobias, Gamarra-Toledo, García-Olaechea, Pulgarín, & Seddon, 2011). Outcomes of this study suggest that following the initial evolution of a mutually ornamented ancestor, sex differences and dimorphism in a broad suite of visual and acoustic traits were most likely created and maintained in barn swallows through countervailing selection on females, rather than on males (Kraaijeveld, Kraaijeveld-Smit, & Komdeur, 2007).

Conclusions

With accumulating evidence for mutual mate choice and mutual ornamentation as the ancestral state across diverse taxa (Edward & Chapman, 2011; Hofmann et al., 2008; Kraaijeveld et al., 2007; Odom et al., 2014), we offer a method for testing whether selection on females may promote sex differences to a much greater extent than is currently appreciated. 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). There are growing accounts of temperate zone female birds that sing for brief periods during the breeding season, particularly during settlement of breeding territories (Hathcock & Benedict, 2018; Krieg & Getty, 2016; Taff, Littrell, & Freeman-Gallant, 2012). Thus, more focused research effort during this 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.

Although there is increasing evidence that female signalling traits can evolve as fast as or faster than male traits across species (Dunn et al., 2015; Johnson, Jordan Price, & Pruett-Jones, 2013; Price, 2019), we have few studies within species to guide our understanding of the processes driving these patterns. In contrast to the vast literature on within-species geographical divergence in male signals (McLean & Stuart-Fox, 2014; Podos & Warren, 2007; Slabbekoorn & Smith, 2002; Velásquez, 2014), there are few studies for females in the visual modality (McCoy, Harmon, Baird, & Fox, 1997; McLean & Stuart-Fox, 2014; Obara, Ozawa, & Fukano, 2008; Roulin, 2003; Tuomaala, Kaitala, & Rutowski, 2012) and fewer still in the acoustic modality (Graham, Heath, Walter, Mark, & Mennill, 2017, 2018; Mennill & Rogers, 2006; Odom & Mennill, 2012). Yet, it is worth noting that three studies (in butterflies and birds) that considered intraspecific geographical variation in both sexes found greater signal divergence in females than in males across populations (Graham et al., 2018; Mennill & Rogers, 2006; Tuomaala et al., 2012). The implications of female signal evolution and male mate choice/recognition as well as their impact on speciation have received little attention in the literature, with a few relevant studies focusing primarily on fish and insects (Chung et al., 2014; Edward & Chapman, 2011; Jiggins, Estrada, & Rodrigues, 2004; Roberts & Mendelson, 2017, 2020). Targeted study of both sexes in concert across taxa is necessary to gain a more holistic understanding of how signals evolve within and among populations and how these trait changes feed into larger ecological, evolutionary and ecoevolutionary processes (Bonduriansky, 2011; Cole & Endler, 2015; Endler & Basolo, 1998; Fryxell, Weiler, Kinnison, & Palkovacs, 2019).

Funding

This work was supported by an American Ornithological Society: Joseph Grinnell Award and a U.S. National Science Foundation (NSF) Graduate Research Fellowship to M.R.W. K.J.O. was supported by NSF Postdoctoral Research Fellowship in Biology (grant no. 1612861).

Author Contributions

M.R.W. collected data, performed analysis and wrote the initial manuscript draft; K.J.O. measured and extracted acoustic metrics from audio files; L.B. undertook major restructuring of the manuscript; and R.J.S. provided project administration and resources for carrying out the work. All authors contributed significantly to manuscript writing and revision.

Acknowledgments

We thank the Macaulay Library for the loan of audio equipment to M.R.W. in 2011, and numerous undergraduate and graduate students in the Safran Lab who helped collect the mountain of phenotypic and fitness data that enabled this study. We also thank the many site owners who allowed us to traipse around their property at odd hours of the day (in particular, the Colorado Horse Rescue). Finally, we thank Megan Barkdull for help with sound analysis and Joey Hubbard for sage counsel on colour analysis.

Supplementary Material

Supplementary material associated with this article is available at https://doi.org/10.1016/j.anbehav.2020.07.018.

References

Akre, K. L., Farris, H. E., Lea, A. M., Page, R. A., & Ryan, M. J. (2011). Signal perception in frogs and bats and the evolution of mating signals. *Science, 333*(6043), 751–752.

Akre, K. L., & Johnsen, S. (2014). Psychophysics and the evolution of behavior. *Trends in Ecology & Evolution, 29*(5), 291–300.

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, NJ: Princeton University Press.

Badyaev, A. V., & Hill, G. E. (2003). Avian sexual dichromatism in relation to phylogeny and ecology. *Annual Review of Ecology, Evolution and Systematics, 34*(1), 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, 54*(3), 987–997.

Bell, R. C., & Zamudio, K. R. (2012). Sexual dichromatism in frogs: Natural selection, sexual selection and unexpected diversity. *Proceedings of the Royal Society B: Biological Sciences, 279*(1748), 4687–4693.

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

Bioacoustics Research Program. (2011). *Raven Pro: Interactive sound analysis software* Version 1.5. http://www.birds.cornell.edu/raven.

Blankers, T., Oh, K. P., & Shaw, K. L. (2019). Parallel genomic architecture underlies repeated sexual signal divergence in Hawaiian crickets. *Proceedings of the Royal Society B: Biological Sciences, 286*(1912), 20191479.

Bonduriansky, R. (2011). Sexual selection and conflict as engines of ecological diversification. *American Naturalist, 178*(6), 729–745.

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. (2020). Barn swallow - *Hirundo rustica. Birds of the world*. In P. G. Rodewald (Ed.), *Birds of the world*. Ithaca, NY: Cornell Lab of Ornithology. https://doi.org/10.2173/bow.barswa.01 Version 1.0. .

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*, Article 13. 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. *Auk: Ornithological Advances, 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, 52*(4), 1219.

Byers, J., Hebets, E., & Podos, J. (2010). Female mate choice based upon male motor performance. *Animal Behaviour, 79*(4), 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*, Article 119. 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, U.K.: 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.

Chung, H., Loehlin, D. W., Dufour, H. D., Vaccarro, K., Millar, J. G., & Carroll, S. B. (2014). A single gene affects both ecological divergence and mate choice in *Drosophila*. *Science, 343*(6175), 1148–1151.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: L. Erlbaum.

Cole, G. L., & Endler, J. A. (2015). Variable environmental effects on a multicomponent sexually selected trait. *American Naturalist, 185*(4), 452–468.

Collins, S. A., de Kort, S. R., Pérez-Tris, J., & Tellería, J. L. (2009). Migration strategy and divergent sexual selection on bird song. *Proceedings of the Royal Society B: Biological Sciences, 276*(1656), 585–590.

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, E. (1801). *Zoonomia; Or, the laws of organic life*. London, U.K.: J. Johnson.

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

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, 114*(10), 955–964.

Dunn, P. O., Armenta, J. K., & Whittingham, L. A. (2015). Natural and sexual selection act on different axes of variation in avian plumage color. *Science Advances, 1*(2), Article e1400155. https://doi.org/10.1126/sciadv.1400155

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

Eikenaar, C., Whitham, M., Komdeur, J., van der Velde, M., & Moore, I. T. (2011). Testosterone, plumage colouration and extra-pair paternity in male North-American barn swallows. *PloS One, 6*(8), Article e23288.

Endler, J. A., & Basolo, A. L. (1998). Sensory ecology, receiver biases and sexual selection. *Trends in Ecology & Evolution, 13*(10), 415–420.

Endler, J. A., & Mielke, P. (2005). Comparing entire colour patterns as birds see them. *Biological Journal of the Linnean Society, 86*, 405–431.

Evans, M. R. (1998). Selection on swallow tail streamers. *Nature, 394*(6690), 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.

Fryxell, D. C., Weiler, D. E., Kinnison, M. T., & Palkovacs, E. P. (2019). Eco-evolutionary dynamics of sexual dimorphism. *Trends in Ecology & Evolution, 34*(7), 591–594.

Galeotti, P., Saino, N., Perani, E., Sacchi, R., & Møller, A. R. (2001). Age-related song variation in male barn swallows. *Italian Journal of Zoology, 68*(4), 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., et al. (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, 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, 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., & Bitz, J. H. (2004). Functions of duetting in the tropical boubou, *Laniarius aethiopicus*: Territorial defence and mutual mate guarding. *Animal Behaviour, 68*(1), 193–201. https://doi.org/10.1016/j.anbehav.2003.11.003

Graham, B. A., Heath, D. D., Walter, R. P., Mark, M. M., & Mennill, D. J. (2018). Parallel evolutionary forces influence the evolution of male and female songs in a tropical songbird. *Journal of Evolutionary Biology, 31*(7), 979–994. https://doi.org/10.1111/jeb.13279

Graham, B. A., Sandoval, L., Dabelsteen, T., & Mennill, D. J. (2017). A test of the acoustic adaptation hypothesis in three types of tropical forest: Degradation of male and female rufous-and-white wren songs. *Bioacoustics, 26*(1), 37–61. https://doi.org/10.1080/09524622.2016.1181574

Hare, R. M., & Simmons, L. W. (2019). Sexual selection and its evolutionary consequences in female animals. *Biological Reviews of the Cambridge Philosophical Society, 94*(3), 929–956.

Hasiniaina, A. F., Radespiel, U., Kessler, S. E., Evasoa, M. R., Rasoloharijaona, S., Randrianambinina, B., et al. (2020). Evolutionary significance of the variation in acoustic communication of a cryptic nocturnal primate radiation (*Microcebus* spp.). *Ecology and Evolution, 10*(8), 3784–3797.

Hathcock, T. J., & Benedict, L. (2018). Conspecific challenges provoke female canyon wrens to sing but not to duet. *Behavioral Ecology and Sociobiology, 72*, Article 196. 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 of the Royal Society B: Biological Sciences, 283*(1826), Article 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 differences 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 (*Icterus* spp.). *Auk: Ornithological Advances, 125*(4), 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. *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. *Auk: Ornithological Advances, 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.

Jiggins, C. D., Estrada, C., & Rodrigues, A. (2004). Mimicry and the evolution of premating isolation in *Heliconius melpomene* Linnaeus. *Journal of Evolutionary Biology, 17*(3), 680–691.

Johnson, A. E., Jordan Price, J., & Pruett-Jones, S. (2013). Different modes of evolution in males and females generate dichromatism in fairy-wrens (Maluridae). *Ecology and Evolution, 3*(9), 3030–3046.

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.

Keen, S. C., Odom, K. J., Kohn, G., Wright, T., Webster, M. S., & Araya-Salas, M. (2020). *A machine-learning approach for quantifying element diversity in animal vocalizations*. Manuscript submitted for publication.

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. *Evolutionary Ecology, 26*(5), 1101–1118. https://doi.org/10.1007/s10682-012-9563-5

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

Kraaijeveld, K., Kraaijeveld-Smit, F. J. L., & Komdeur, J. (2007). The evolution of mutual ornamentation. *Animal Behaviour, 74*(4), 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. (2016). 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. https://CRAN.R-project.org/package=rsample.

Lande, R. (1980). Sexual differences, sexual selection, and adaptation in polygenic characters. *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.

Lee, K.-H., Shaner, P.-J. L., Lin, Y.-P., & Lin, S.-M. (2016). Geographic variation in advertisement calls of a microhylid frog: Testing the role of drift and ecology. *Ecology and Evolution, 6*(10), 3289–3298.

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

Magoolagan, L., Mawby, P. J., Whitehead, F. A., & Sharp, S. P. (2019). The structure and context of male and female song in white-throated dippers. *Journal of Ornithology, 160*(1), 195–205. https://doi.org/10.1007/s10336-018-1599-z

McCoy, J. K., Harmon, H. J., Baird, T. A., & Fox, S. F. (1997). Geographic variation in sexual dichromatism in the collared lizard, *Crotaphytus collaris* (Sauria: Crotaphytidae). *Copeia, 1997*(3), 565–571. https://doi.org/10.2307/1447560

McGraw, K. J., Safran, R. J., & Wakamatsu, K. (2005). How feather colour reflects its melanin content. *Functional Ecology, 19*(5), 816–821. https://doi.org/10.1111/j.1365-2435.2005.01032.x

McLean, C. A., & Stuart-Fox, D. (2014). Geographic variation in animal colour polymorphisms and its role in speciation. *Biological Reviews of the Cambridge Philosophical Society, 89*(4), 860–873.

Mennill, D. J., & Rogers, A. C. (2006). Whip it good! Geographic consistency in male songs and variability in female songs of the duetting eastern whipbird *Psophodes olivaceus*. *Journal of Avian Biology, 37*(1), 93–100. https://doi.org/10.1111/j.2006.0908-8857.03548.x

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 differences, but not contemporary evolution. *Journal of Evolutionary Biology, 29*(11), 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. *Nature, 332*(6165), 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., et al. (2019). The causal relationship between sexual selection and sexual size differences in marine gastropods. *Animal Behaviour, 148*, 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: Neuroethology, Sensory, Neural, and Behavioral Physiology, 188*(11–12), 1003–1014.

Obara, Y., Ozawa, G., & Fukano, Y. (2008). Geographic variation in ultraviolet reflectance of the wings of the female cabbage butterfly, *Pieris rapae*. *Zoological Science, 25*(11), 1106–1110.

Odom, K. J., & Benedict, L. (2018). A call to document female bird songs: Applications for diverse fields. *Auk: Ornithological Advances, 135*(2), 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*, Article 3379.

Odom, K. J., & Mennill, D. J. (2012). Inconsistent geographic variation in the calls and duets of barred owls (*Strix varia*) across an area of genetic introgression. *Auk: Ornithological Advances, 129*(3), 387–398.

Orme, D., Freckleton, R., Thomas, G., Petzoldt, T., Fritz, S., Isaac, N., et al. (2018). *caper: Comparative analyses of phylogenetics and evolution in R* Version 1.0.1. https://CRAN.R-project.org/package=caper.

Pap, P. L., Fülöp, A., Adamkova, M., Cepak, J., Michalkova, R., Safran, R. J., et al. (2019). Selection on multiple sexual signals in two central and eastern European populations of the barn swallow. *Ecology and Evolution, 9*(19), 11277–11287. https://doi.org/10.1002/ece3.5629

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

Pieplow, N. (2017). *Peterson field guide to bird sounds of eastern North America*. Boston, MA: Houghton Mifflin Harcourt.

Podos, J., & Warren, P. S. (2007). The evolution of geographic variation in birdsong. *Advances in the Study of Behavior, 37*, 403–458. https://doi.org/10.1016/s0065-3454(07)37009-5

Price, T. D. (1984). The evolution of sexual size differences in Darwin's finches. *American Naturalist, 123*(4), 500–518.

Price, J. J. (2015). Rethinking our assumptions about the evolution of bird song and other sexually dimorphic signals. *Frontiers in Ecology and Evolution, 3*, Article 40. https://doi.org/10.3389/fevo.2015.00040

Price, J. J. (2019). Sex differences in song and plumage color do not evolve through sexual selection alone: New insights from recent research. *Journal of Ornithology, 160*(4), 1213–1219. https://doi.org/10.1007/s10336-019-01681-8

Price, J. J., & Eaton, M. D. (2014). Reconstructing the evolution of sexual dichromatism: Current color diversity does not reflect past rates of male and female change. *Evolution, 68*(7), 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 of the Royal Society B: Biological Sciences, 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, 11*(3), 247–251.

R Core Team. (2019). *R: A language and environment for statistical computing* Version 3.6.0. 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*, Article 62. https://doi.org/10.3389/fevo.2016.00062

Riebel, K., Hall, M. L., & Langmore, N. E. (2005). 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.

Roberts, N. S., & Mendelson, T. C. (2017). Male mate choice contributes to behavioural isolation in sexually dimorphic fish with traditional sex roles. *Animal Behaviour, 130*, 1–7.

Roberts, N. S., & Mendelson, T. C. (2020). *Identifying female phenotypes that promote behavioral isolation in a sexually dimorphic species of fish (Etheostoma zonale)*. bioRxiv, Article 224. biorxiv;2020.04.20.051714v1.

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, 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 of the Royal Society B: Biological Sciences, 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. *Behavioral Ecology, 30*(6), 1653–1661. https://doi.org/10.1093/beheco/arz130

Rosvall, K. A. (2011). Intrasexual competition in females: Evidence for sexual selection? *Behavioral Ecology, 22*(6), 1131–1140.

Roulin, A. (2003). Geographic variation in sexual differences in the barn owl *Tyto alba*: A role for direct selection or genetic correlation? *Journal of Avian Biology, 34*(3), 251–258. https://doi.org/10.1034/j.1600-048x.2003.03022.x

Ryan, M. J., Akre, K. L., Baugh, A. T., Bernal, X. E., Lea, A. M., Leslie, C., et al. (2019). Nineteen years of consistently positive and strong female mate preferences despite individual variation. *American Naturalist, 194*(2), 125–134.

Safran, R. J., Adelman, J. S., McGraw, K. J., & Hau, M. (2008). Sexual signal exaggeration affects physiological state in male barn swallows. *Current Biology, 18*(11), R461–R462. https://doi.org/10.1016/j.cub.2008.03.031

Safran, R. J., & McGraw, K. (2004). Plumage coloration, not length or symmetry of tail-streamers, is a sexually selected trait in North American barn swallows. *Behavioral Ecology, 15*(3), 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., et al. (2016). The maintenance of phenotypic divergence through sexual selection: An experimental study in barn swallows *Hirundo rustica*. *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.

Shultz, A.llison J., & Burns, K. J. (2017). The role of sexual and natural selection in shaping patterns of sexual dichromatism in the largest family of songbirds (Aves: Thraupidae). *Evolution, 71*(4), 1061–1074.

Sibley, D. A. (2014). *The Sibley guide to birds*. New York, NY: A. A. Knopf.

Slabbekoorn, H., & Smith, T. B. (2002). Bird song, ecology and speciation. *Philosophical Transactions of the Royal Society B: Biological Sciences, 357*(1420), 493–503. https://doi.org/10.1098/rstb.2001.1056

Smith, C. C. R., Flaxman, S. M., Scordato, E. S. C., Kane, N. C., Hund, A. K., Sheta, B. M., et al. (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), Article 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. *Journal of Evolutionary Biology, 24*(7), 1517–1532. https://doi.org/10.1111/j.1420-9101.2011.02284.x

Stoddard, M. C., & Prum, R. O. (2008). Evolution of avian plumage color in a tetrahedral color space: A phylogenetic analysis of new world buntings. *American Naturalist, 171*, 755–776.

Stokes, D., & Stokes, L. (2010). *The Stokes field guide to the birds of North America*. London, U.K.: Little, Brown Book Group.

Svensson, E. I., & Waller, J. T. (2013). Ecology and sexual selection: Evolution of wing pigmentation in calopterygid damselflies in relation to latitude, sexual differences, and speciation. *American Naturalist, 182*(5), E174–E195.

Taff, C. C., Littrell, K. A., & Freeman-Gallant, C. R. (2012). Female song in the common yellowthroat. *Wilson Journal of Ornithology, 124*(2), 370–374. https://doi.org/10.1676/11-182.1

Tobias, J. A., Gamarra-Toledo, V., García-Olaechea, D., Pulgarín, P. C., & Seddon, N. (2011). Year-round resource defence and the evolution of male and female song in suboscine birds: Social armaments are mutual ornaments. *Journal of Evolutionary Biology, 24*(10), 2118–2138. https://doi.org/10.1111/j.1420-9101.2011.02345.x

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

Tuomaala, M., Kaitala, A., & Rutowski, R. L. (2012). Females show greater changes in wing colour with latitude than males in the green-veined white butterfly, *Pieris napi* (Lepidoptera: Pieridae). *Biological Journal of the Linnean Society, 107*(4), 899–909. https://doi.org/10.1111/j.1095-8312.2012.01996.x

Velásquez, N. A. (2014). Geographic variation in acoustic communication in anurans and its neuroethological implications. *Journal of Physiology, 108*(2–3), 167–173.

Webb, W. H., Brunton, D. H., Aguirre, J. D., Thomas, D. B., Valcu, M., & Dale, J. (2016). Female song occurs in songbirds with more elaborate female coloration and reduced sexual dichromatism. *Frontiers in Ecology and Evolution, 4*, Article 22. https://doi.org/10.3389/fevo.2016.00022

Wiens, J. J. (2001). Widespread loss of sexually selected traits: How the peacock lost its spots. *Trends in Ecology & Evolution, 16*(9), 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., et al. (2018). Global song divergence in barn swallows (*Hirundo rustica*): Exploring the roles of genetic, geographical and climatic distance in sympatry and allopatry. *Biological Journal of the Linnean Society, 123*(4), 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 of the Royal Society B: Biological Sciences, 282*(1816), Article 20151574.

Yasukawa, K. (1989). The costs and benefits of a vocal signal: The nest-associated 'chit' of the female red-winged blackbird, *Agelaius phoeniceus*. *Animal Behaviour, 38*(5), 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.

Appendix

Table A1
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 (kHz)
Acoustic	Frequency range	Average range of raw frequencies measured across song (kHz)
Acoustic	Mean peak frequency	Average frequency with the highest energy from the mean frequency spectrum (kHz)
Acoustic	Element count	Number of song elements (syllables) per song
Acoustic	Element duration	Average duration of elements per individual (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 (s)
Acoustic	Gap duration	Average length of intervals between elements (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 (kHz)
Acoustic	CV_entropy	Coefficient of variation for spectrographic entropy
Acoustic	CV_dom frequency range	Coefficient of variation for the range of the dominant frequency measured across the song (kHz)
Acoustic	CV_frequency range	Coefficient of variation for raw frequencies measured across song (kHz)
Acoustic	CV_mean peak frequency	Coefficient of variation for frequency with the highest energy from the mean frequency spectrum (kHz)
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 (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 (s)
Acoustic	CV_gap duration	Coefficient of variation for length of intervals between elements (s)
Acoustic	CV_element diversity	Coefficient of variation for the average minimum convex polygon area surrounding the element level acoustic space for each song
Acoustic	CV_frequency IQR	Coefficient of variation for interquartile range for frequency (kHz)

Table A1 (*continued*)

Trait type	Trait	Definition
Visual	Tail length	Maximum length of two longest tail feathers (streamers), each of which was the average of 3 repeated measures (mm)
Visual	Brightness	The average percentage 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 colour range (600–700 nm); higher values darker; measured for throat, breast, belly and vent

Table A2

Bootstrap analysis of 35 acoustic and visual traits in male and female barn swallows, showing sample sizes, means and standard deviations for each trait, sex differences (i.e. Cohen's *d*) and sex difference 95% confidence intervals (CI) for 10 000 bootstraps

Trait	Female		Male		Sex difference	95% CI
	N	Mean (SD)	*N*	Mean (SD)		
Belly brightness[a]	**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[a]	**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[a]	**12**	**1.53e-05 (8.21e-06)**	**15**	**3.44e-05 (8.21e-06)**	**−3.11**	**(−1.88, −1.53)**
Element duration[a]	**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[a]	**12**	**0.831 (0.0121)**	**15**	**0.792 (0.0121)**	**3.27**	**(1.54, 1.85)**
Frequency IQR[a]	**12**	**0.916 (0.189)**	**15**	**1.04 (0.189)**	**−0.816**	**(−1.45, −0.0324)**
Frequency range[a]	**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[a]	**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[a]	**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[a]	**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)

Sample sizes for coefficients of variation are lower for females, as three individuals only had a single song recorded. Traits shown in bold had sex difference confidence intervals that did not overlap zero.

[a] Traits selected for the final analysis as shown in Fig. 4, after eliminating the less biologically intuitive trait for redundant trait pairs with >|0.7| correlations.

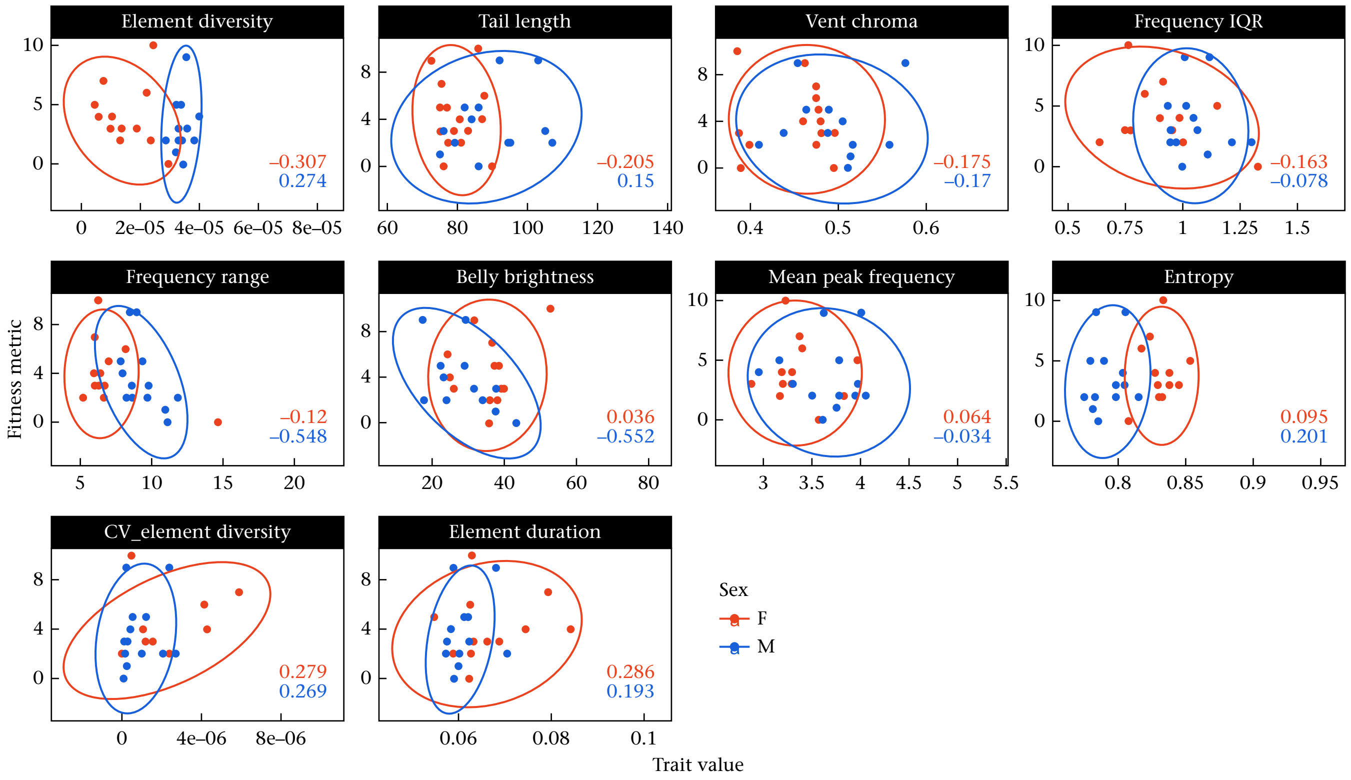


Figure A1. Spearman's rank correlations between robustly dimorphic traits shown in Fig. 4 and a relevant fitness metric — our measure of contemporary selection. The fitness metric was seasonal within-pair genetic paternity for males and seasonal fledging success for females. Orange 95% confidence ellipses 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.

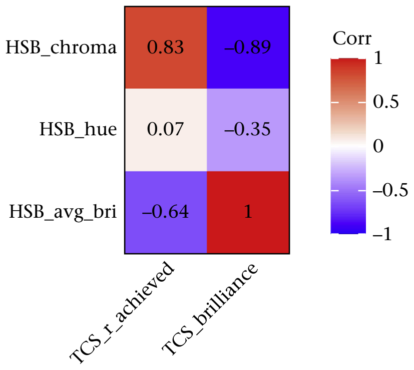


Figure A2. Spearman's rank correlations between our reported hue–saturation–brightness (HSB) colour measures and tetrahedral colour space (TCS) colour metrics. All measures are for the 27 individuals (both sexes) in our data set with complete feather colour measures. Only the breast patch is shown, since correlations are similar across patches in this subspecies.