# **Title:** Correlated evolution of distinct signals associated with increased social selection in female white-shouldered fairywrens

**Running title:** Correlated evolution of female signals

John Anthony Jones1,2, Karan J Odom3,4, Ian R Hoppe5, Doka Nason6, Serena Ketaloya6, and Jordan Karubian1

1 Department of Ecology and Evolutionary Biology, Tulane University, New Orleans, USA 70118

2 Corresponding email: johnajones91@gmail.com

3 Cornell Lab of Ornithology and Department of Neurobiology and Behavior, Cornell University, Ithaca, NY USA

4 Current address: Department of Psychology, University of Maryland, College Park, MD USA

5 School of Natural Resources, University of Nebraska, Lincoln, USA 68583

6 Porotona Village, Milne Bay Province, Papua New Guinea

# **Abstract –**

Conspicuous female signals have recently received substantial scientific attention, but the degree to which female-specific selection drives their evolution remains unclear. Species that express female-specific phenotypic variation among populations represent a useful opportunity to address this knowledge gap. White-shouldered fairywrens (*Malurus alboscapulatus*) are tropical songbirds with a well-resolved phylogeny where female, but not male, coloration varies allopatrically across subspecies. We explored how four distinct signaling modalities, each putatively associated with increased social selection, are expressed in two populations that vary in competitive pressure on females. Females in a derived subspecies (*M. a. moretoni*) have evolved more ornamented plumage, more complex vocalizations, and shorter tails (a signal of social dominance) relative to an ancestral subspecies (*M. a. lorentzi*) with drab females. Moreover, in response to simulated territorial intrusions broadcasting female song, female *M. a. moretoni* are more aggressive and more coordinated with their mate in both movement and vocalizations. These results suggest that correlated phenotypic shifts in female color, morphology, song complexity, and behavior may have occurred in response to changes in social selection, consistent with the idea that female-specific selection has driven the evolution of these signals.

*Keywords*: fairywren, female ornamentation, female song, social selection

# **Introduction –**

Improving our knowledge of sexual dimorphism and phenotypic diversity depends on understanding signal evolution in both sexes. For this reason, better resolving the factors underlying variation in female signals is a core goal among contemporary evolutionary ecologists (Clutton-Brock 2007, Tobias et al. 2012, Odom et al. 2014, Webb et al. 2016, Doutrelant et al. 2020). One key question concerns the degree to which plumage-based visual signals and vocal signals (e.g., song) are under direct selection in females versus being effectively dragged along by active selection on males (Darwin 1871, Lande 1980). While the non-adaptive hypothesis has received some support (Poissant et al. 2010; Kraaijeveld 2014), there are cases in which female visual and vocal signals are different from or more elaborate than that of males (e.g., Heinsohn et al. 2005, Brunton and Li 2005, Illes and Yunes-Jimenez 2008). Mounting evidence suggests that female ornamentation can evolve independently from that of males (Odom et al. 2014, Dale et al. 2015, Wilkins et al. 2020).

Animals often simultaneously use multiple signal modalities, either to reinforce a signal’s message or to convey distinct messages to different receivers (Møller and Pomiankowski 1993). Studies that concurrently evaluate multimodal signal evolution in the context of a known phylogeny are rare but are imperative to understand female trait evolution (e.g., Gomes et al. 2017, Hasegawa et al. 2017). For example, evidence from two or more populations with known phylogenetic history and different female signal states may provide insights into evolutionary relationships between female-specific selection and signal expression. Female-specific selection may often take the form of social selection, a process including competition for mates (i.e., sexual selection) as well as ecological resources (West-Eberhard 1979, 1983). Although social selection is increasingly used as a framework to evaluate the evolutionary significance of female ornamentation (Lyon and Montgomerie 2012, Tobias et al. 2012), studies linking female signal evolution to changes in strength of social selection are rare (Doutrelant et al. 2020).

Correlated evolution of distinct signal types in relation to shifts in social selection pressures would support the hypothesis that female-specific selection drives female signal expression. Here, we define correlated signal evolution as the emergence of a suite of signaling traits occurring in the same direction (e.g., evolution of more complex song along with more ornamented plumage), without implying that they emerge at the same time or via the same physiological mechanism. Rubenstein and Lovette (2009) used a qualitatively similar approach to show that female plumage ornamentation of different species of starling (Sturnidae) increases in parallel with intrasexual competition for mates. However, additional examples, particularly involving evolution of multiple signal types among different populations of a single species, are lacking.

White-shouldered fairywrens (*Malurus alboscapulatus*) of tropical New Guinea present a study system in which female signals vary between closely-related populations and ancestral vs. derived states of female plumage phenotype can be inferred with confidence. This species is unusual, in that female plumage varies geographically, while male coloration remains unchanged (Rowley and Russell 1997, Karubian 2013, Enbody et al. 2019, Fig. 1A). Female ornamentation in this species is derived: subspecies with ornamented females have evolved from unornamented female ancestors independent of any transition in male ornamentation (Driskell et al. 2011, Johnson et al. 2013, Karubian 2013, ED Enbody unpubl. data). Both sexes sing during territory defense (Rowley and Russel 1997), but detailed analyses of song are lacking. Both sexes from a subspecies in which females are ornamented (*M. a. moretoni*; Enbody et al. 2018; Fig. 1A) are more aggressive and have higher circulating testosterone than a subspecies with unornamented females (*M. a. lorentzi*). Among *Malurus* fairywrens, shorter tails appear to be associated with social dominance (Swaddle et al. 2000, Karubian et al. 2009). Past work showed that unadjusted tail lengths of *M. a. moretoni* and *M. a.* *lorentzi* females did not vary statistically (Enbody et al. 2019) but did not account for body size differences. Finally, *M. a. lorentzi* commonly interacts with neighbors outside of the pair in a social context, whereas *M. a. moretoni* experiences limited extra-territory interactions (JA Jones, unpubl. data), suggesting that female-specific selection pressures experienced via competition are greater in *M. a. moretoni* (see also below).

We combined a field-based simulated territorial intrusion (STI) experiment, song analysis, and morphological measures to test the hypothesis that increased female plumage ornamentation is accompanied by parallel transitions in acoustic, behavioral, and morphological traits, consistent with increased social selection. We used STIs to explore the behavioral response of territorial residents to intruders of varying female plumage and song phenotype (Fig. 1). We first predicted that female *M. a. moretoni* would respond to STIs by both local and foreign phenotypes (1) more aggressively at the individual level and (2) with greater coordination at the pair level than would *M. a. lorentzi* individuals and pairs, consistent with Enbody et al. (2018). If these traits evolve in concert, we also predicted increased song complexity, but shorter body-size adjusted tail length in female *M. a. moretoni*, in tandem with greater plumage ornamentation. Finally, if our initial predictions are supported, we predict that the combination of the *M. a. moretoni* plumage and song will elicit the strongest aggressive response in both populations. Support for these predictions would suggest that increases in signal complexity and female ornamentation are associated with an increase in joint territory defense. Moreover, this would suggest that these distinct signaling modalities have undergone correlated evolutionary transitions in a direction consistent with increased social selection.

# **Methods –**

We studied two populations of white-shouldered fairywrens from May-July 2018: *M. a. moretoni* (in which both males and females are ornamented) in Garuahi Village, Milne Bay Province (Lat: -10.2187, Long: 150.4843) and *M. a. lorentzi* (in which females are unornamented; Fig. 1); in Obo Village, Western Province, Papua New Guinea (-7.6047, 141.3090). We captured fairywrens via mist-nets, banded each bird with a unique combination of colored leg bands for individual identification, and measured tail and tarsus lengths (±0.01 mm).

#### Behavioral assay protocol:

Enbody et al. (2018) previously used a pair-level simulated territorial intrusion (STI) with paired male and female mounts of both ornamented and unornamented phenotypes. In both cases, the mounts were coupled with audio duet (male and female) exemplars of the resident (local) song and were not presented with the other subspecies’ song. The authors found that mount phenotype (i.e., ornamented versus unornamented) did not affect how pairs responded. We built upon this earlier study by assessing behavioral responses to varying female plumage and song phenotype stimuli, without any male stimuli. This design simulates a female intruder on a mated pair’s territory, and thus the biological interpretation of this intrusion may be that the simulated intruder is attempting to usurp the female. This distinction is noteworthy because whereas an intruding male and female pair may represent an equal threat to both sexes, a lone female could be perceived as more of a threat to resident females than resident males (e.g., Mennill 2006). As our treatments were a combination of female plumage and song exemplars only (i.e., no male stimuli), expectations for how the mated male may respond to these intrusions is less straightforward. Males may not perceive a solo female song or mount as a sufficient threat to warrant an intense aggressive response as sexes often compete for separate resources that affect different fitness parameters.

Our experimental design follows that of Greig et al. (2015), who compared male behavior and ornamentation in two allopatric subspecies of red-backed fairywrens (*M. melanocephalus*). Free-flying pairs were repeatedly assayed (without replacement) with one of with five possible female plumage and song combinations in randomized order with ~3 days (range: 2-5 days) between trials (Fig. 1C): (1) *M. a. lorentzi* song and plumage mount, (2) *M. a. moretoni* song and *M. a. lorentzi* plumage, (3) *M. a. moretoni* song and plumage, (4) *M. a. lorentzi* song and *M. a. moretoni* mount. Additionally, we presented (5) an emperor fairywren (*M. cyanocephalus*) female mount paired with local (with respect to population) white-shouldered fairywren song to serve as a heterospecific plumage control. We did not use an acoustic control, as neither subspecies responds to heterospecific playback. Our aim was too present pairs with each of the five treatments; we did not include assay trials if one member of the mated pair did not respond, regardless of sex.

We broadcasted song exemplars previously recorded (during other field seasons; no playback exemplars were recorded from individuals within this study from this year) and broadcast from the mount location using an Ultimate Ears Roll 2 speaker (Irvine, CA, USA). Songs were randomly chosen, after excluding song from the focal individual and neighboring territories. Apart from the emperor fairywren phenotype, mount exemplars (n = 4 of each phenotype; 12 total) were the same three-dimensional painted bird models used in Enbody et al. (2018).

Free-living pairs were detected without audio duet playback in most cases (exposing birds to duets prior to behavioral assays did not significantly influence the response (p >0.30), after which mounts were placed ~1.5 m off the ground and we retreated ~25 m (range: 20-40 m) and minimized exposure. For ~5 min of playback and 5 min of silence (post playback), we recorded latency to respond and approach (song response is not discernible without visual observation), proportion of time within 5 m of the mount, and the average distance to mount (calculated as the average time spent within each categorical distance class: <0.5 m, 0.5–5 m, 5–10 m, 10–15 m, and >15 m) separately for the free-flying male and female. Additionally, we noted individual flybys within 2 m of the mount as well as songs sung with and without the mate. To quantify a pair’s degree of coordination in response to treatment, we noted the latency between the male and female responses (i.e., shorter latency is associated with more coordinated behaviors) as well as the proportion of time the pair spent together (within 1 m of each other) throughout the trial. We recorded other pair coordination behaviors, including allopreening, duetting, and leapfrogs (where one bird jumps over another on the same perch) as rates (events per min). We noted whether a male attempted to court the mount (i.e., visual displays: puff-shoulder display, display flights, and petal carries), but found there was no relationship among treatments nor between subspecies and display rate (all p >0.32), and thus exclude these results from the current study.

## *Song analysis*:

Female songs were recorded using a Marantz PMD 661 Mk II (96kHz sampling rate, 24-bit depth; D&M Professional, Itasca, IL) with a Sennheiser ME66 shotgun microphone and K6 power module (Sennheiser Electronic Corporation, Old Lyme, CT). For each population, we created playback exemplars from five individual females recorded in 2017. Each exemplar consisted of a single female song repeated at 10 sec intervals for ~5 min total (exact time varied; trial length was accounted for to calculate behavior rates). We used Audacity to filter out noise below 500 Hz and standardize amplitude.

We calculated element diversity and several other metrics representing song complexity to compare overall song structure of male and female songs recorded in 2018 (both subspecies) and 2019 (*M. a. moretoni* only). We first standardized to a sample rate of 44.1 kHz and bit depth of 16 and then used Raven Sound Analysis Software v1.6 (Center for Conservation Bioacoustics 2019) to select every 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. Measurements in Raven were made using a Hanning window with a 512 FFT and 90% overlap for a time resolution of 1.161 msec, and a frequency resolution of 86.1 Hz. From these selections, we extracted robust (energy-based) time and frequency measurements of each element in Raven. We then transferred the selections into R using the *Rraven* package (Araya-Salas 2017), where we extracted additional acoustic parameters via the *warbleR* package (Araya-Salas and Smith-Vidaurre 2017). We removed highly correlated acoustic variables (r > |0.95|), resulting in 24 acoustic variables extracted for each element, including measures of frequency, frequency bandwidth, frequency modulation, time, duration, and entropy (Table S1). These element-level metrics were used to estimate element diversity from a 2-D acoustic trait space (Keen et al. *in press*). We also used the song\_param function in *warbleR* to extract song-level metrics, resulting in the following final set of metrics calculated for the entirety of each song: (1) song duration, (2) mean element duration, (3) mean peak frequency, (4) frequency range (calculated as the difference between the highest 95% and the lowest 5% frequency values of all elements in each song), and (5) element diversity. Here, 5% and 95% frequencies were calculated in Raven and are the frequencies that divide the selection into two frequency intervals containing 5% and 95% of the energy in the selection (Hz), and therefore represent minimum and maximum frequency, respectively. These five song-level metrics were used in subsequent statistical analyses to compare overall song structure between the subspecies.

To calculate element diversity, we created an element-level acoustic space via an unsupervised random forest that included every element-level acoustic parameter in R package *randomForest* (Liaw and Wiener 2002). The random forest analysis was run with the following specifications: 10000 trees, minimal node size of 1, Gini impurity index as split rule, five randomly sampled variables at each split and out-of-bag proximity. This process created a proximity matrix that was transformed into a set of five vectors using classic multidimensional scaling (MDS) via the cmdscale function in the *stats* R package. We used the MDS vector to create a 2-D acoustic space containing all elements of all songs. The area that the element encompasses is indicative of its diversity, with larger areas being more diverse. We extracted 95% minimum convex polygon of the areas defined by the elements for each song using the function mcp in the R package *adehabitatHR* (Calenge 2015). This method was ground-truthed using datasets of known element diversity by Keen et al. (*in press*).

## *Statistical analysis*:

We reanalyzed a subset of morphological descriptive data available from Enbody et al. (2019) to test whether tail length varies between subspecies while controlling for body size and sex. We removed outliers (defined as individuals with tail lengths >1.5 times the interquartile range away from the mean), controlled for body size by regressing tail length against tarsus length, and performed an independent samples t-test to test for sex-specific differences in residual tail length between subspecies.

We explored response to simulated intruders at the level of the individual (i.e., male or female response independent of one another) and pair (i.e., the joint, coordinated response) separately using two principal components analyses (PCA). In the Individual-PC, we included the following individual-level metrics: latency to respond, flyby rate, solo and duet song rate, time within 5 m of the mount, and average distance from the mount. For the Pair-PC, we included the following pair-level metrics: the difference in time between male and female response to the mount (denoted as latency lag), the proportion of time spent together versus apart during the trial, and the rates of allopreening, leapfrogging, and duetting. Duets were included in both PCs to account for the fact that some individuals may have intended to sing individually but were subsequently joined in song by the partner. For all behavioral PCs, we normalized each response variable by log+1 transformation followed by centering and scaling prior to running the PCA (Filardi and Smith 2008, Uy et al. 2009).

We first assessed the effects of stimulus treatment type, subspecies, and sex of the responding bird on the top principal components using linear mixed effects models using the *lme4* package in R (Bates et al. 2015; R Core Team 2020). Individual ID (individual analysis only) was nested within pair ID as random intercepts in our models, as well as mount ID (12 available mount phenotypes, 4 per treatment) and female song stimulus ID (10 available song recordings, 5 per subspecies) to reduce the effect of pseudoreplication (see Kroodsma et al. 2001). Covariates including pair breeding stage (i.e., breeding versus non-breeding), time of day of the behavior assay, the order in which each treatment was given to a pair, and whether or not another male/juvenile approached during the assay did not significantly predict the behavioral response, and thus were not included in the final model. Residuals of the full model did not violate assumptions of normality nor homoscedasticity. The significance of each model was evaluated using a χ2 test in the *car* package in R (Fox and Weisberg 2011).

To explore whether song elaboration varies at the sex- and population-level, we created a reduced set of song-level acoustic variables via a PCA on a correlation matrix of our acoustic variables and extracted the first three principal components for further analyses. We then ran three separate linear mixed effects models for each principal component, with those components as dependent variables and with sex and subspecies as fixed effects. We included individual ID, population, and year recorded as random intercepts in each model. We additionally tested for a significant interaction between subspecies and sex; upon finding the interaction was not statistically significant, we ran our models without the interaction term.

## **Ethics statement:**

Our study was carried out in strict accordance with the guidelines established by the Tulane University Institutional Animal Care and Use Committee (#0395R2). We minimized handling time for each individual in order to reduce physical stress and harm.

# **Results –**

#### Morphological comparison:

Female *M. a. moretoni* tails are shorter on average than those of *M. a. lorentzi* when accounting for body size (t = 5.32, df = 199, p < 0.001), whereas male body-size adjusted tail length does not significantly vary between populations (t = 0.11, df = 206, p = 0.91).

#### Individual response to simulated territorial intrusion:

We excluded trials when a bird did not respond to the stimulus, resulting in 91 *M. a. lorentzi* and 102 *M. a. moretoni* female responses and 97 *M. a. lorentzi* and 102 *M. a. moretoni* male responses. The first three PCs explained 76.9% of the variation in individual responses (Table 1). Higher scores for Individual-PC1 are associated with individuals that are more aggressive, characterized predominately by a faster response to the stimulus, aggressive posturing (i.e., birds spending more time close to the mount), and more frequent flybys. Higher scores for Individual-PC2 indicate individuals that sing more solo songs and tend to fly by the mount more often. Individual-PC3 appears to be associated with individuals that favor threats rather than direct aggression, such that higher scores correspond to quicker response times and greater singing rates, but with a tendency to remain >5 m from the mount.

There were no significant interactions among treatment, sex, or subspecies, and thus we excluded these interaction terms from the final models. The lack of interaction between treatment and subspecies suggests that the two subspecies respond to treatments in a similar fashion. Thus, each subspecies’ response is determined by whether the stimulus treatment is local or foreign, rather than presented phenotype was ornamented or not. Treatment predicted Individual-PC1 and Individual-PC3 (Table 2); individuals of both sexes are more aggressive when exposed to treatments that contain a local song (with respect to population; all p <0.01, Fig. 2A, Table S2). Additionally, *M. a. moretoni* respond overall more aggressively than *M. a. lorentzi* (higher PC1 scores; Fig. 2B). Individual-PC3 varies between sexes, such that males are overall more responsive to simulated territorial intrusions on this axis. Finally, Individual-PC2 is not predicted by treatment or subspecies, but we found a near significant trend that females tend to have lower PC2 scores, indicative of fewer flybys of the exemplar coupled with increased duetting (rather than individual solo songs) when responding to playback (p = 0.06, Table 2).

#### Pair response to simulated territorial intrusion:

We excluded pair responses in which only one member of the pair responded to the stimulus, resulting in 90 trials completed in *M. a. lorentzi* and 102 in *M. a. moretoni*. For these trials, the first two PCs explain ~60% of the variation in pair responses to simulated territorial intrusion (Table 1). Higher scores for Pair-PC1 indicate a greater degree of pair coordination. This was characterized by more frequent duets and allopreening and a greater proportion of time spent together. Pair-PC2 appears to be associated with birds that spend a greater proportion of time together with a smaller latency lag (i.e., a more synchronous response to the stimulus), but also perform allopreening and leapfrog behaviors less frequently. We interpret higher Pair-PC2 scores as indicating a more coordinated and aggressive response.

We found no significant interaction between treatment and subspecies and did not include these interaction terms in the final model. Treatment predicted Pair-PC1, but not Pair-PC2 (Table 2). Similar to Individual-PC1, the differences among treatments are driven by the presence of the local phenotype’s song: pair coordination is higher in simulated territorial intrusions when a local song (with respect to the focal population) is broadcasted, regardless of color mount presented (all p < 0.01, Fig. 3A, Table S2). Subspecies, but not treatment, predicted Pair-PC2; *M. a. moretoni* are more coordinated in response to simulated intrusions than are *M. a. lorentzi* pairs (Fig. 3B, Table 2).

## *Song*:

We recorded 45 female songs (n = 11 individuals) and 26 male songs (n=14) in one population of *M. a. lorentzi* and 36 female (9) and 32 male songs (11) between two populations of *M. a. moretoni*. We present the top three principal components (Table 3), explaining 69.7% of the total acoustic variation. The first component explains 34.2% of the total variation and captures element structure and variation. High scores for this component indicate songs with longer and more diverse elements separated from one another by comparatively short periods of silence. Highly-scoring songs also show a greater frequency range at the element level, although the mean peak frequency of the entire song is lower. Song-PC2 is driven predominately by the frequency range at the song level, with higher scores corresponding to songs with a greater frequency range. Finally, negative Song-PC3 scores are associated with longer songs, while positive scores tend to correspond to songs with longer elements with greater frequency ranges despite these not loading as highly as duration.

We found no sex differences across the models but did find a significant difference between subspecies in Song-PC1 and -PC3 (Table 4**)**. Both Song-PCs are higher in *M. a. moretoni*: their songs contain longer, more diverse, and more complex elements than those sung by *M. a. lorentzi*, whose songs are longer and occur at a greater peak frequency on average than their ornamented counterparts.

# **Discussion –**

Many female passerines exhibit elaborate plumage and sing complex songs, but only recently have we begun to rigorously explore the evolutionary antecedents of these traits. We investigated four distinct phenotypic traits in two subspecies of *M. alboscapulatus*, testing whether those traits diverged in the same direction (i.e., increasing in ornamentation and complexity) in association with an increase in social selective pressure. Each of the four signals we investigated (female tail length, song complexity, aggression, and pair coordination) are putatively involved in social interactions all four traits exhibit a correlated evolutionary transition that parallels an increase in female plumage ornamentation from a drab ancestral state to a derived, ornamented state. The most parsimonious explanation for this finding is that direct selection is operating on female fairywrens to promote several, potentially reinforcing signal types: female-specific divergence in four independent phenotypic traits commonly associated with social selection, either via sexual selection (i.e., intrasexual competition for mates, intersexual mate choice) or competition (e.g., over ecological, non-reproductive resources), is unlikely to be happenstance. However, at this time we are unable to determine the order in which these traits arose evolutionarily, nor whether there is a common physiological mechanism (e.g., testosterone) underlying these transitions.

These results provide an opportunity to better understand how population-level differences in social selection pressure may be associated with female phenotypic divergence. Several key differences in the life history of our two study populations likely contribute to differing social selection pressures on females. First, the population with dull female plumage (*M. a. lorentzi*) experiences pronounced seasonal (monsoonal) variation whereas *M. a. moretoni* occurs in a more stable climate (Enbody et al. 2019) which in turn is generally associated with year-round territoriality, reduced sexual promiscuity, and more equal effort between the sexes (Stutchbury and Morton 2001, Tobias et al. 2016). Moreover, female *M. a.* *lorentzi* exhibit lower concentrations of circulating testosterone relative to *M. a.* *moretoni* (Enbody et al. 2018), consistent with a lower degree of intrasexual selection pressure.

Increases in social cohesion within mated pairs often arise in the context of greater competitive pressures for ecological resources (Griffith 2019). *M. a. moretoni* songs consist of longer, more diverse elements than those of *M. a. lorentzi*, which may be perceived as higher quality (e.g., Ballentine 2006, Searcy and Nowicki 2006). Although both of our focal fairywren populations duet, we found evidence for greater coordination among pairs in *M. a. moretoni*. The occurrence of duetting is strongly associated with year-round territoriality and strong social bonds, and playback studies support that duets in many species are used in territory defense and maintaining a functioning partnership (Wachtmeister 2001, Dahlin and Benedict 2014, Tobias et al. 2016). Although alternative explanations for population-level acoustic differences exists, we believe these are unlikely the key determinant of phenotypic divergence observed here. For example, while aspects of the ecological environment may influence the efficacy of acoustic signals (Endler 1993), both subspecies sing predominately in open grasslands where transmission efficacy is similar (JA Jones pers. obs.).

In contrast to our predictions, local vocal signals elicited a uniformly stronger response than did plumage coloration, suggesting a prioritization of acoustic over visual signals in this system. Both subspecies recognized and responded aggressively to the song phenotype of the opposite population, suggesting elements of the acoustic structure are retained between populations and are used in species recognition contexts. Like red-backed fairywren males (Greig et al. 2015), our focal species discriminated between song types, but not color, at the subspecies level. Several non-mutually exclusive explanations for this result are possible. The capacity for song to elicit aggressive behaviors may be strong enough to swamp other types of signals, such that variation in those signals is functionally irrelevant in aggressive contexts involving song. Alternatively, individuals may gauge the quality of the territory intruder and appropriately respond based on their own coloration. For example, male red-backed fairywrens are less aggressive towards potential rivals that have lower-quality plumage (Karubian et al. 2008). Of course, female plumage may be assessed by potential mates rather than by rivals (e.g., Thys et al. 2020). These results are not an indication that visual signals are irrelevant in female competitive contexts, but they do suggest that signaling function of female *M. a. moretoni* plumage ornaments warrants further study.

Selection should favor a multidimensional integrated phenotype that better equips individuals to respond to changing selective pressures. Phenotypic integration occurs when multiple life-history characteristics, including morphological, physiological, and/or behavioral traits covary in a manner that is more adaptive than if the traits were decoupled (see reviews in Pigliucci and Preston 2004, McGlothlin and Ketterson, 2008). Testosterone is a possible candidate hormone that plays a role in facilitating phenotypic integration (reviewed in Lipshutz et al. 2019). Higher concentrations of testosterone in *M. a. moretoni* females compared to *M. a.* *lorentzi* females is consistent with a role in coordinating expression of both plumage and behavior (Enbody et al. 2018). Moreover, Boersma et al. (2020) were able to experimentally induce the white shoulder ornament in the population with unornamented females (*M. a. lorentzi*) via an exogenous testosterone implant. However, ornament expression in unornamented females is not a naturally occurring phenotype, and it remains unclear if androgens have an influence on ornament expression in populations with naturally occurring female ornamentation. It is additionally unclear what role testosterone plays on tail length or acoustic performance (e.g., Dittrich et al. 2014 and references therein). These represent an exciting avenue for future work in this system to determine the degree to which testosterone underlies one or more of these traits.

More broadly, although we often characterize female signals by the degree to which they resemble those of males, in many instances selection may often operate on both sexes independently (Rosvall 2011, Cain and Rosvall 2014, Dale et al. 2015). As such, a more nuanced focus on how female-specific selection occurs and directly influences the evolution of female signals per se is appropriate. This study, which highlights this evidence for female-specific selection by detailing how multiple signaling phenotypes exhibit correlated evolutionary transitions in the context of increasing social selection pressures, represents one step in this direction.

**Acknowledgements:**

This research would not be possible without the Baivapupu Clan of Obo Village, Western Province, for granting us access to their land. We thank Morrison Olesai and Kiplin Saiga for assistance in the field. We additionally thank the Milne Bay and Western Province provincial governments for permits and permissions as well as the local level landowners for granting access to each site and privately owned land. The National Research Institute assisted in acquiring our country-level research permits (#99902100765) and visas. Finally, we are grateful to members of the Karubian Lab at Tulane University, as well as Jordan Boersma and Erik Enbody for providing thoughtful comments on earlier versions of this manuscript. This research was funded by grants from the National Science Foundation (IOS-1354133) awarded to JK, the American Ornithological Society (to JAJ) and the Department of Ecology and Evolutionary Biology of Tulane University (to JAJ). Funders have had no influence on the content of the submitted manuscript nor do they require approval of the final manuscript to be published.

**Author Contributions**:

JAJ and JK designed the experiment. JAJ, IRH, DN, and SK carried out the experiment. JAJ, IRH, and KJO carried out the statistical analyses and interpretation of results. JAJ wrote the first draft and each coauthor subsequently provided substantial input towards the final version.

**Data Accessibility Statement:**

Data and R script files are available publicly on Figshare at doi:[10.6084/m9.figshare.14137406](https://doi.org/10.6084/m9.figshare.14137406) and doi:[10.6084/m9.figshare.14135966](https://doi.org/10.6084/m9.figshare.14135966).

# **Literature Cited –**

Araya-Salas, M. 2017. Rraven: connecting R and Raven bioacoustic software. R package 1.0.0.

Araya-Salas, M., and G. Smith-Vidaurre. 2017. warbleR: an r package to streamline analysis of animal acoustic signals. Methods Ecol. Evol. 8:184–191.

Ballentine, B. 2006. Morphological adaptation influences the evolution of a mating signal. Evolution. 60:1936–1944.

Bates, D., M. Mächler, B. M. Bolker, and S. C. Walker. 2015. Fitting Linear Mixed-Effects Models Using lme4. J. Stat. Softw. 67:1–48.

Boersma, J., E. D. Enbody, J. A. Jones, D. Nason, E. Lopez-Contreras, J. Karubian, and H. Schwabl. 2020. Testosterone induces plumage ornamentation followed by enhanced territoriality in a female songbird. Behav. Ecol., doi: 10.1093/beheco/araa077.

Brunton, D. H., and X. Li. 2006. The song structure and seasonal patterns of vocal behavior of male and female bellbirds (*Anthornis melanura*). J. Ethol. 24:17–25.

Cain, K. E., and K. A. Rosvall. 2014. Next steps for understanding the selective relevance of female-female competition. Front. Ecol. Evol. 2:2012–2014.

Calenge, C. 2015. adehabitatHS: analysis of habitat selection by animals. R Package. version 0.3.12.

Center for Conservation Bioacoustics. 2019. Raven Pro: Interactive sound analysis software (Version 1.6.1) [Computer software]. Ithaca, NY: The Cornell Lab of Ornithology. Available from http://ravensoundsoftware.com/.

Clutton-Brock, T. 2007. Sexual selection in males and females. Science. 318:1882–1885.

Dale, J., C. Dey, K. Delhey, B. Kempenaers, and M. Valcu. 2015. The effects of life-history and social selection on male and female plumage coloration. Nature 527:367–370.

Dahlin, C. R., and L. Benedict. 2014. Angry birds need not apply: A perspective on the flexible form and multifunctionality of avian vocal duets. Ethology 120:1–10.

Darwin, C. 1871. The descent of man and selection in relation to sex. Murray, London.

Dittrich, F., C. Ramenda, D. Grillitsch, C. Frankl-Vilches, M. C. Ko, M. Hertel, W. Goymann, A. ter Maat, and M. Gahr. 2014. Regulatory mechanisms of testosterone-stimulated song in the sensorimotor nucleus HVC of female songbirds. BMC Neurosci. 15:1–16.

Doutrelant, C., A. Fargevieille, and A. Grégoire. 2020. Evolution of female coloration: What have we learned from birds in general and blue tits in particular. Adv. Study Behav. 52:123–202.

Driskell, A. C., J. A. Norman, S. Pruett-Jones, E. Mangall, S. Sonsthagen, and L. Christidis. 2011. A multigene phylogeny examining evolutionary and ecological relationships in the Australo-Papuan wrens of the subfamily Malurinae (Aves). Mol. Phylogenet. Evol. 60:480–485.

Enbody, E. D., J. Boersma, H. Schwabl, and J. Karubian. 2018. Female ornamentation is associated with elevated aggression and testosterone in a tropical songbird. Behav. Ecol. 29:1056–1066.

Enbody, E. D., J. Boersma, J. A. Jones, M. W. H. Chatfield, S. Ketaloya, D. Nason, D. T. Baldassarre, J. Hazlehurst, O. Gowen, H. Schwabl, and J. Karubian. 2019. Social organisation and breeding biology of the White-shouldered Fairywren (*Malurus alboscapulatus*). Emu 119:274–285.

Endler, J. A. 1993. Some general comments on the evolution and design of animal communication systems. Philos. Trans. R. Soc. London, B 340:215–225.

Filardi, C. E., and C. E. Smith. 2008. Social selection and geographic variation in two monarch flycatchers from the Solomon Islands. Condor 110:24–34.

Fox, J., and S. Weisberg. 2011. An {R} companion to applied regression, Second Edition. SAGE Publications Thousand Oaks (CA).

Gomes, A. C. R., C. Funghi, M. Soma, M. D. Sorenson, and G. C. Cardoso. 2017. Multimodal signalling in estrildid finches: song, dance and colour are associated with different ecological and life-history traits. J. Evol. Biol. 30:1336–1346.

Greig, E. I., D. T. Baldassarre, and M. S. Webster. 2015. Differential rates of phenotypic introgression are associated with male behavioral responses to multiple signals. Evolution. 69:2602–2612.

Griffith, S. C. 2019. Cooperation and coordination in socially monogamous birds: Moving away from a focus on sexual conflict. Front. Ecol. Evol. 7:1–15.

Heinsohn, R., S. Legge, and J. A. Endler. 2005. Extreme reversed sexual dichromatism in a bird without sex role reversal. Science. 309:617–619.

Illes, A. E., and L. Yunes-Jimenez. 2012. A female songbird out-sings male conspecifics during simulated territorial intrusions. Proc. R. Soc. B Biol. Sci. 276:981–986.

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

Karubian, J. 2013. Female ornamentation in *Malurus* fairy-wrens: A hidden evolutionary gem for understanding female perspectives on social and sexual selection. Emu 113:248–258.

Karubian, J., T. S. Sillett, and M. S. Webster. 2008. The effects of delayed plumage maturation on aggression and survival in male red-backed fairy-wrens. Behav. Ecol. 19:508–516.

Karubian, J., J. P. Swaddle, C. W. Varian-Ramos, and M. S. Webster. 2009. The relative importance of male tail length and nuptial plumage on social dominance and mate choice in the red-backed fairy-wren Malurus melanocephalus: Evidence for the multiple receiver hypothesis. J. Avian Biol. 40:559–568.

Keen, S.C., Odom, K.J., Webster, M.S., Kohn, G.M., Wright, T.F., and Araya-Salas, M. A machine learning approach for classifying and quantifying acoustic diversity. Methods in Ecology and Evolution*. In press.*

Kraaijeveld, K. 2014. Reversible trait loss: The genetic architecture of female ornaments. Annu. Rev. Ecol. Evol. Syst. 45:159–177.

Kroodsma, D. E., B. E. Byers, E. Goodale, S. Johnson, and W.-C. Liu. 2001. Pseudoreplication in playback experiments, revisited a decade later. Anim. Behav. 61:1029–1033.

Lande, R. 1980. Sexual dimorphism, sexual selection, and adaption in polygenic characters. Evolution. 34:292–305.

Liaw, A., and M. Wiener. 2002. Classification and regression by randomForest. R News 2:18–22.

Lipshutz, S. E., E. M. George, A. B. Bentz, and K. A. Rosvall. 2019. Evaluating testosterone as a phenotypic integrator: From tissues to individuals to species. Mol. Cell. Endocrinol. 496:110531.

Lyon, B. E., and R. Montgomerie. 2012. Sexual selection is a form of social selection. Philos. Trans. R. Soc. B Biol. Sci. 367:2266–2273

McGlothlin, J. W., and E. D. Ketterson. 2008. Hormone-mediated suites as adaptations and evolutionary constraints. Philos. Trans. R. Soc. B Biol. Sci. 363:1611–1620.

Møller, A. P., and A. Pomiankowski. 1993. Why have birds got multiple sexual ornaments? Behav. Ecol. Sociobiol. 32:167–176.

Odom, K. J., M. L. Hall, K. Riebel, K. E. Omland, and N. E. Langmore. 2014. Female song is widespread and ancestral in songbirds. Nat. Commun. 5:3379.

Pigliucci, M., and K. Preston. 2004. Phenotypic integration: Studying the ecology and evolution of complex phenotypes. Oxford University Press, New York.

Poissant, J., A. J. Wilson, and D. W. Coltman. 2010. Sex-specific genetic variance and the evolution of sexual dimorphism: A systematic review of cross-sex genetic correlations. Evolution. 64:97–107.

R Core Team. 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: <https://www.R-project.org/>.

Rowley, I., and E. Russell. 1997. Bird families of the world: Fairy-wrens and Grasswrens. Oxford University Press, Oxford, United Kingdom.

Rosvall, K. A. 2011. Intrasexual competition in females: Evidence for sexual selection? Behav. Ecol. 22:1131–1140.

Rubenstein, D. R., and I. J. Lovette. 2009. Reproductive skew and selection on female ornamentation in social species. Nature 462:786–789.

Searcy, W. A., and S. Nowicki. 2006. Signal interception and the use of soft song in aggressive interactions. Ethology 112:865–872.

Stutchbury, B. J. M., and E. S. Morton. 2001. Behavioral ecology of tropical birds. Academic Press, London, United Kingdom.

Swaddle, J. P., J. Karubian, and S. Pruett-Jones. 2000. A novel evolutionary pattern of reversed sexual dimorphism in fairy wrens: Implications for sexual selection. Behav. Ecol. 11:345–349.

Thys, B., R. Pinxten, and M. Eens. 2020. Does the tie fit the female? Melanin-based colouration, aggressive personality and reproductive investment in female great tits (*Parus major*). Behav. Ecol. Sociobiol. 74:43.

Tobias, J. A., C. Sheard, N. Seddon, A. Meade, A. J. Cotton, and S. Nakagawa. 2016. Territoriality, social bonds, and the evolution of communal signaling in birds. Front. Ecol. Evol. 4:74.

Tobias, J. A., R. Montgomerie, and B. E. Lyon. 2012. The evolution of female ornaments and weaponry: social selection, sexual selection and ecological competition. Philos. Trans. R. Soc. B 367:2274–2293.

Uy, J. A. C., R. G. Moyle, and C. E. Filardi. 2009. Plumage and song differences mediate species recognition between incipient flycatcher species of the Solomon Islands. Evolution. 63:153–164.

Wachtmeister, C. A. 2001. Display in monogamous pairs: A review of empirical data and evolutionary explanations. Anim. Behav. 61:861–868.

Webb, W. H., D. H. Brunton, J. D. Aguirre, D. B. Thomas, M. Valcu, and J. Dale. 2016. Female song occurs in songbirds with more elaborate female coloration and reduced sexual dichromatism. Front. Ecol. Evol. 4:22.

West-Eberhard, M. J. 1979. Sexual selection, social competition, and evolution. Proc. Am. Philos. Soc. 123:222–234.

West-Eberhard, M. J. 1983. Sexual selection, social competition, and speciation. Q. Rev. Biol. 58:155–183.

Wilkins, M. R., K. J. Odom, L. Benedict, and R. J. Safran. 2020. Analysis of female song provides insight into the evolution of sex differences in a widely studied songbird. Anim. Behav. 168:69–82. Elsevier Ltd.

# **Tables**

Table 1. Loading scores for the principal component analysis exploring (1) individual and (2) paired response to simulated territorial intrusions.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Individual-PC1** | **Individual-PC2** | **Individual-PC3** | **Pair-PC1** | **Pair-PC2** |
| Eigenvectors (SD) | 1.594 | 1.045 | 0.992 | 1.413 | 0.999 |
| Proportion of Variance | 0.423 | 0.182 | 0.164 | 0.399 | 0.200 |
|  |  |  |  |  |  |
| Latency | -0.381 | 0.086 | -0.555 |  |  |
| Flyby (min-1) | 0.364 | 0.309 | -0.132 |  |  |
| Time within 5m of mount | 0.553 | 0.077 | -0.327 |  |  |
| Avg. distance from mount | -0.580 | -0.042 | 0.225 |  |  |
| Solo songs (min-1) | 0.010 | 0.771 | 0.515 |  |  |
| Duet (min-1) | 0.283 | -0.544 | 0.503 | 0.591 | 0.118 |
| Latency Lag |  |  |  | -0.340 | 0.680 |
| Allopreen (min-1) |  |  |  | 0.466 | 0.440 |
| Leapfrog (min-1) |  |  |  | 0.395 | 0.354 |
| Proportion of time together | |  |  | 0.404 | -0.452 |

*Individual-PC: Latency: delay in response to playback stimuli. Flyby: the rate at which individuals flew past 2m of the mount. Pair-PC: Latency lag: difference in time between the first responses to playback stimulus by each member of the focal pair; proportion of time: time spent physically together and behaving in a coordinated fashion versus time spent physically apart. Note: duets are a component of both PCs.*

Table 2. Summary statistics detailing the effects of treatment, subspecies, and sex on the behavioral response of individuals at both the individual (top) and pair (bottom) levels.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Individual response to simulated territorial intrusion** | | | | | | |
|  | **Fixed Effects** | **Estimate** | **Std error** | **t-value** | **F** | **df** | **P** |
| Individual-PC1 | Treatment | -0.71 | 0.17 | 25.8 | 21 | 4, 41.485 | **<0.001** |
|  | Subspecies | 0.17 | 0.17 | 26.93 | 10.1 | 1, 41.527 | **0.003** |
|  | Sex | 0.35 | 0.17 | 24.66 | 0.35 | 1, 329.795 | 0.555 |
|  | *Random effects:* | *Variance* |  |  |  |  |  |
|  | Bird ID/Pair ID | <0.001 |  |  |  |  |  |
|  | Pair ID | 1.16 |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
| Individual-PC2 | Treatment | -0.09 | 0.15 | 23.79 | 1.56 | 4, 37.247 | 0.207 |
|  | Subspecies | 0.22 | 0.15 | 24.67 | 0.83 | 1, 39.644 | 0.369 |
|  | Sex | 0.07 | 0.15 | 22.62 | 3.83 | 1, 38.079 | *0.058* |
|  | *Random effects:* | *Variance* |  |  |  |  |  |
|  | Bird ID/Pair ID | 0.09 |  |  |  |  |  |
|  | Pair ID | 0.21 |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
| Individual-PC3 | Treatment | -0.40 | 0.16 | 20.02 | 4.2 | 4, 39.245 | **0.006** |
|  | Subspecies | -0.39 | 0.16 | 20.45 | 0.87 | 1, 38.272 | 0.357 |
|  | Sex | -0.08 | 0.16 | 19.75 | 6.9 | 1, 39.912 | **0.012** |
|  | *Random effects:* | *Variance* |  |  |  |  |  |
|  | Bird ID/Pair ID | 0.01 |  |  |  |  |  |
|  | Pair ID | 0.34 |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  | **Pair response to simulated territorial intrusion** | | | | | | |
|  | **Fixed effects** | **Estimate** | **Std error** | **t-value** | **F** | **df** | **P** |
| Pair-PC1 | Treatment | -0.65 | 0.244 | 34.56 | 5.6 | 4, 45.18 | **0.001** |
|  | Subspecies | 0.05 | 0.222 | 32.71 | 1.98 | 1, 41.27 | 0.167 |
|  | *Random effects* | *Variance* |  |  |  |  |  |
|  | Pair ID | 0.74 |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
| Pair-PC2 | Treatment | -0.65 | 0.244 | 34.56 | 1.08 | 4, 46.86 | 0.375 |
|  | Subspecies | 0.05 | 0.222 | 32.71 | 11.7 | 1, 40.36 | **0.001** |
|  | *Random effects* | *Variance* |  |  |  |  |  |
|  | Pair ID | 0.08 |  |  |  |  |  |

Table 3. PCA eigenvalues, proportion of variation explained, and loading factors for the top three components of acoustic variation.

|  |  |  |  |
| --- | --- | --- | --- |
| **Acoustic parameters** | **Song-PC1** | **Song-PC2** | **Song-PC3** |
| Eigenvalues | 1.547 | 1.202 | 1.020 |
| Proportion of variance | 0.342 | 0.206 | 0.148 |
| Song duration | 0.287 | 0.297 | -0.576 |
| Mean element duration | 0.451 | 0.056 | 0.472 |
| Mean peak frequency | -0.414 | 0.393 | 0.389 |
| Frequency range (song-level) | -0.156 | 0.682 | -0.083 |
| Mean frequency range (element-level) | 0.425 | 0.281 | 0.479 |
| Gap duration | -0.436 | 0.249 | 0.025 |
| Element diversity | 0.384 | 0.385 | -0.240 |

*Song and element duration are defined as the units of time for the length of the song and mean element length. Frequency range at the song and element levels are calculated from the highest 95% and the lowest 5% frequency value for all elements in each song. Gap duration is the mean unit of time separating elements within a song. Element diversity is calculated as a 95% minimum complex polygon surrounding song structure at the element-level acoustic space.*

Table 4. Model estimates for how acoustic parameters vary between sex and subspecies.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | **Estimate** | **Std error** | **t-value** | **F** | **df** | **P** |
| Song-PC1 | Sex | -1.034 | 0.264 | 4.124 | 0.281 | 1, 49.35 | 0.598 |
|  | Subspecies | 0.127 | 0.24 | 49.350 | 36.69 | 1, 3.772 | **0.005** |
| Song-PC2 | Sex | 0.056 | 0.221 | 39.849 | 0.576 | 1, 49.505 | 0.452 |
|  | Subspecies | -0.206 | 0.272 | 49.505 | 0.237 | 1, 45.484 | 0.628 |
| Song-PC3 | Sex | 0.149 | 0.173 | 38.918 | 1.346 | 1, 48.943 | 0.252 |
|  | Subspecies | 0.244 | 0.211 | 48.943 | 6.31 | 1, 44.32 | **0.016** |

# **Figures**

Figure 1. Visual representation of natural variation in signaling modalities between populations of white-shouldered fairywren. (A) *Malurus alboscapulatus lorentzi* of Western Provence, Papua New Guinea (shaded circle; left) and *M. a. moretoni* of Milne Bay Province (open circle; right). (B) Representative example sonogram of *M. a. lorentzi* (top; brown cartoon) and *M. a. moretoni* (bottom, black cartoon) female song. (C) Experimental design of the current study (illustrated by picture of each subspecies to represent the mount plumage presented and the cartoon fairywren to represent song exemplars as given in (B). Free flying fairywrens were presented with each treatment in randomized order: (1) *M. a. lorentzi* song and plumage mount, (2) *M. a. moretoni* song and *M. a. lorentzi* plumage, (3) *M. a. moretoni* song and plumage, (4) *M. a. lorentzi* song and *M. a. moretoni* mount, and (5) *M. cyanocephalus* plumage mount (to serve as a control) paired with the local phenotype’s song phenotype.

Figure 2. White-shouldered fairywren behavioral response (PC1) to simulated territorial intrusions with respect to treatment and subspecies. Bars represent mean ± standard error.

Figure 3. White-shouldered fairywren pair coordinated response to simulated territorial intrusions with respect to (A) treatment for Pair-PC1 and (B) subspecies for Pair-PC2. Bars represent mean ± standard error.

Figure 1.

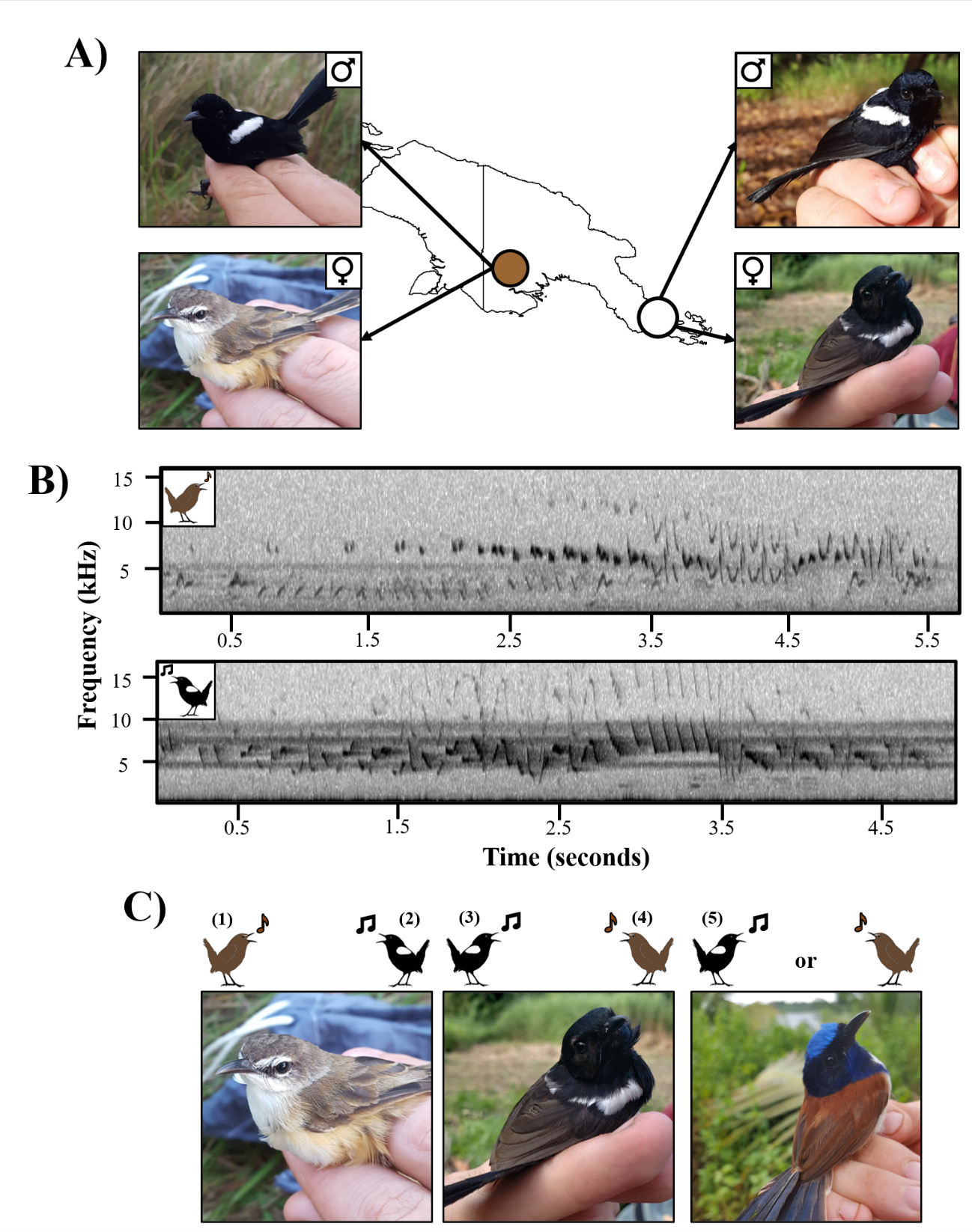


Figure 2.

Chart, box and whisker chart

Description automatically generated

Figure 3.

Chart, box and whisker chart

Description automatically generated