Behavioral Ecology
The official journal of the
ISBE
International Society for Behavioral Ecology

Behavioral Ecology (2018), 00(00), 1–11. doi:10.1093/beheco/ary079

Original Article

Female ornamentation is associated with elevated aggression and testosterone in a tropical songbird

Erik D. Enbody[a], Jordan Boersma[b], Hubert Schwabl[b], and Jordan Karubian[a]
[a]Department of Ecology and Evolutionary Biology, Tulane University, 400 Lindy Boggs, New Orleans, LA 70118, USA and [b]School of Biological Sciences, Washington State University, PO Box 644236, Pullman, WA 99164, USA

Received 7 December 2017; revised 26 April 2018; editorial decision 5 May 2018; accepted 9 May 2018.

In males, testosterone plays a key role in ornament production and linking ornamentation with reproductive behaviors and other traits to produce an integrated phenotype. Less is known about whether females couple testosterone, ornamentation, and aggressive behaviors to achieve female-specific combinations of traits. Ornamentation in females may be the result of correlated expression with male ornamentation, or female traits could arise as the result of sex-specific selection pressures. Resolving between these alternatives is necessary to understand the degree to which selection acts on female traits. The White-shouldered Fairywren (*Malurus alboscapulatus*) provides a useful context to address these questions because populations vary in degree of female ornamentation, a derived trait, whereas male ornamentation is constant across both populations. We found that ornamented females have higher levels of circulating testosterone and respond more aggressively to experimental territorial intrusions than do unornamented females. These findings are consistent with the idea that, among female White-shouldered Fairywrens, testosterone may mechanistically link plumage and behavioral traits to produce an integrated competitive phenotype, as has been reported for males of closely related species. In contrast, circulating testosterone in males did not differ significantly between populations. More broadly, our findings are consistent with ongoing selection on the mechanisms underlying female ornaments, likely via social selection.

Key words: **aggression, bird, female ornamentation, integrated phenotype, passerine, testosterone.**

INTRODUCTION

Sex-based differences in visual signals, which are widespread in nature, have long captured the attention of naturalists and evolutionary ecologists and their study has provided important insight into natural and sexual selection. A comprehensive understanding of sex-based differences depends upon understanding the underlying mechanisms of production and patterns of expression in both sexes. Yet, until recently, most studies of ornamentation have focused on males, despite females of many taxa exhibiting ornamentation (Amundsen 2000; Heinsohn 2005; Clutton-Brock and Huchard 2013). Some have proposed that female ornamentation is a nonadaptive byproduct of selection favoring ornaments in males (Darwin 1871; Lande 1980) and, consistent with this perspective, female ornaments often closely resemble those of conspecific males (e.g. Amundsen and Parn 2006; Dale et al. 2015). However, female ornaments are sometimes distinct from those of males (e.g. Emlen et al. 2005; Simmons and Emlen 2008; Weiss et al. 2009), and comparative phylogenetic studies have revealed frequent evolutionary transitions between ornamented and unornamented states in females (e.g. Burns 1998; Hofmann et al. 2008; Johnson et al. 2013; Shultz and Burns 2017). These patterns are inconsistent with ornaments in females being selectively neutral byproducts of selection in males (Dale et al. 2015) and suggest that they may evolve independently of males and be the subject of sex-specific selection pressures. At present, however, there is little direct evidence for selection for female ornamentation.

Sex steroids play a key mechanistic role in sex-specific trait expression (Hau 2001), for females as well as males (Staub and De Beer 1997; Kimball and Ligon 1999; Muck and Goymann 2011). Among males, the androgen testosterone is often associated with both individual (Roberts et al. 2009) and interspecific variation in plumage ornaments (Dijkstra et al. 2012; Rosvall et al. 2016) and is the subject of tradeoffs between survival and reproductive investment. Similarly, in females, testosterone can be associated with

---

Address correspondence to E.D. Enbody. E-mail: erik.enbody@gmail.com.

singing (Kriner and Schwabl 1991), aggression (Sandell 2007), nonreproductive physiological processes (Zysling et al. 2006), and fitness (Veiga and Polo 2008; Cain and Ketterson 2012). However, testosterone may have different regulatory roles in females than in males (Goymann and Wingfield 2014) and, in some cases, testosterone levels can also be correlated across the sexes without having a measurable function in females (e.g. Ketterson et al. 2005; Møller et al. 2005; Goymann and Wingfield 2014).

Importantly, testosterone has the potential to link ornament expression with other morphological or behavioral traits to produce an integrated phenotype (Ketterson and Nolan 1992; Cox et al. 2017), in which suites of functionally and mechanistically related traits are coupled in an adaptive manner (Pigliucci 2003). For example, it is well established that changes in circulating testosterone levels regulate the expression of both aggressive and reproductive behavior (Wingfield et al. 1990) and across vertebrates male ornaments are often testosterone regulated (e.g. fish: Fernald 1976, reptiles: Cox et al. 2008, and birds: Lank et al. 1999; Peters et al. 2000; Mougeot et al. 2004; Lindsay et al. 2011). When present in females, ornaments may function in competition over ecological or social resources (West-Eberhard 1979; West-Eberhard 1983), which suggests a potential link between aggression and the adaptive function of female ornaments. Indeed, empirical work has corroborated the idea that aggression in females is often associated with the development of female ornamented traits (e.g. Pryke 2007; Rubenstein and Lovette 2009; Crowhurst et al. 2012) and female birds possess androgen receptors in brain regions associated with aggression (Rosvall et al. 2012). Aggression may also be associated with circulating sex steroids in female birds (Kriner and Schwabl 1991; Pärn et al. 2008; Cain and Ketterson 2012), which may have fitness consequences (Veiga and Polo 2008). Testosterone may become elevated as a response to aggressive encounters (Wingfield et al. 1990; Langmore et al. 2002), but in other species aggression appears to be independent of testosterone levels (Jawor et al. 2006). Similarly, experimental testosterone treatments may stimulate ornament expression in unornamented males of some species (e.g. van Oordt and Junge 1934; Fernald 1976; Mougeot et al. 2004), but not others (Owens and Short 1995), suggesting species-specific effects of testosterone on aggression and ornamentation.

Between the sexes, females may decouple associations between testosterone and traits to produce female-specific phenotypes and behaviors (Rosvall 2013), though the degree to which females are able to do so is unclear. Sexes could use different mechanisms (e.g. hormone target sensitivity, circulation, and conversion to other hormones, or nonendocrine mechanisms, such as gene regulatory mechanisms) to achieve ornamentation (Rosvall et al. 2012), such that ornamentation in females is effectively decoupled from associated behavioral traits observed in males. The evolution of androgen-mediated sexual dimorphism can also be achieved by coupling trait expression to testosterone in males without decoupling these traits from testosterone in females (Cox et al. 2015). Alternatively, the integrated phenotype may remain highly conserved across the sexes so that the same mechanisms involved in the expression of male ornaments and associated behaviors exist in females.

In the current study, we assessed associations between testosterone and territorial aggression between 2 subspecies of the White-shouldered Fairywren (*Malurus alboscapulatus*) that differ in degree of female, but not male, ornamentation (Rowley and Russell 1997; Karubian 2013; Enbody et al. 2017). The White-shouldered Fairywren is sister to the Red-backed Fairywren (*Malurus melanocephalus*) and, although the current study is focused exclusively on White-shouldered Fairywren, some basic information on individual flexibility in ornamentation in Red-backed Fairywren helps to contextualize hypotheses that we present about fixed difference between populations of White-shouldered Fairywrens. Among Red-backed Fairywrens, male plumage ornamentation (Karubian 2002) is associated with increased aggression (Karubian et al. 2008) and both naturally and experimentally elevated circulating testosterone titres (Lindsay et al. 2009; Lindsay et al. 2011). This relationship has led to the expectation that testosterone acts on a mechanistic level to integrate these traits in male Red-backed Fairywrens (Webster et al. 2010). However, recent work has shown that this relationship may vary by age, body condition, and population (Barron et al. 2015; Dowling and Webster 2017; Lantz et al. 2017). The naturally unornamented female Red-backed Fairywrens have lower circulating testosterone than do males (Schwabl et al. 2015), but testosterone-implanted females develop a partial, male-like phenotype (red, but no black coloration; Lindsay et al. 2016). This implies that sex differences in ornamentation of Red-backed Fairywrens are in part mediated by sex differences in circulating levels of testosterone, but the role of testosterone in integration of female ornamentation with behavioral traits is not known.

Resolving the degree to which the integrated ornamented phenotype is conserved between the sexes requires information about the associations between testosterone, ornamentation, and behavioral traits in females, appropriate comparison points with males, and phylogenetic history. Accordingly, the null hypothesis (conservation of mechanisms) predicts that the same combinations of traits (morphological and behavioral) will be regulated via the same mechanism(s) in male and female White-shouldered Fairywrens as in male Red-backed Fairywrens, namely via variation in circulating testosterone levels. Female traits may have developed as a correlated response of selection on males if female hormone levels and behaviors co-vary with males in between-population transitions in ornamentation. Alternatively, female traits may vary independently from males between populations, which would support sex-specific selection pressures. More specifically, if female ornaments are under selection in this system, we would expect that populations of White-shouldered Fairywren with ornamented female plumage show: 1) higher circulating testosterone levels in females, 2) coordinated increases in aggressive behavior in females, and 3) constant levels of testosterone and aggression in males across the 2 populations.

METHODS

Study system and general field methods

Malurus fairywrens are predominantly socially monogamous, though sexually promiscuous passerine birds with an extensive background in life history and behavioral research (Buchanan and Cockburn 2013). The White-shouldered Fairywren (*Malurus alboscapulatus*) of New Guinea forms a monophyletic clade with the sexually dichromatic Red-backed Fairywren and White-winged Fairywren (*Malurus leucopterus*, Driskell et al. 2011) of Australia. Female ornamentation is likely a derived trait in White-shouldered Fairywren populations with female ornamentation, as females are unornamented in both sister taxa (Driskell et al. 2011; Lee et al. 2012; Karubian 2013).

Our study was designed to compare females from 2 populations in Papua New Guinea: one with ornamented females (black-and-white, *M. a. moretoni*) and the other with unornamented females (cryptic brown, *M. a. lorentzi*). Ornamented *M. a. moretoni* females are qualitatively similar to ornamented males, but differ in lacking a dense barbule structure and satin sheen (Enbody et al. 2017). We studied *M. a. moretoni* ("ornamented" population hereafter) in Milne Bay Province (150°30′E, 10°15′S, 0–20 m ASL, Figure 1) and 2) *M. a. lorentzi* ("unornamented" population hereafter) in Western Province (141°19′E, 7°35′S, 10–20 m ASL, Figure 1). Both sites are centered in rural villages surrounded by tropical lowland Savannah. The study site in Milne Bay province includes less contiguous grassland patches and slightly higher (2682 mm/year) and less variable (23% precipitation seasonality, WorldClim) rainfall than the site in Western Province (2339 mm/year, 49% precipitation seasonality, WorldClim).

We monitored populations at the Milne Bay Province site from 2013 to 2016 and the Western Province site during visits in 2014–2016. In each of these periods, we captured most or all adults in an area of approximately 200 ha and monitored their behavior and breeding. As in many other tropical passerines (Stutchbury and Morton 2001), White-shouldered Fairywrens in both populations breed year-round, and males in both populations possess enlarged cloacal protuberances year-round (Enbody ED, Boersma J, unpublished data). Individuals were captured in mist nets by flushing or using playback and banded with a unique combination of plastic color bands and single aluminum band with a unique number provided by the Australian Bird and Bat Banding Scheme. At the time of capture, we collected blood samples (ca. 60–100 μL, mean net to bleeding time = 6 min, range = 1–19 min), which were spun in a centrifuge for 5 min for separation. Plasma was stored in 100% ethanol (Goymann et al. 2007) for analyses.

Androgen radioimmunoassay methods

We used a radioimmunoassay, closely following previously published methods (Lindsay et al. 2009; Lantz et al. 2017), for measuring levels of circulating androgens to approximate testosterone levels for 139 females (105 ornamented, 34 unornamented) and 143 males. The protocol for the present study differed in that plasma samples were stored in ethanol and were only assayed for androgens. Samples ranging from 16.8 to 83.5 μL (mean: 39.49 μL; median: 41.65 μL) were vortexed and centrifuged, and the supernatant was transferred to extrelut columns (MilliporeSigma, Billerica, MA) for extraction of steroids. After extracting with diethyl ether, each sample was redissolved with 125 μL of phosphate-buffered saline with gelatin before being assayed for total androgens. Samples were assayed in single 100 μL aliquots using tritium-labeled testosterone (Perkin Elmer Life Science NET-553, Waltham, MA) and a testosterone antibody (Wien Laboratories, T-3003, Flanders, NJ) that has 100% reactivity with testosterone, 60% with 5α-dihydrotestosterone, 5% with aldosterone, and <15% reactivity with other steroids. Note that we refer to testosterone although our radioimmunoassay cross-reacts with other androgens such as 5α-dihydrotestosterone. We ran samples in singlets rather than duplicates to maximize hormone detection in small sample volumes, following the validated protocol of Lindsay et al. (2009). Samples from both populations and sexes were randomly distributed across 8 separate assays, with a between assay coefficient of variation of 11.46% and a mean within assay coefficient of variation of 7.34% (range = 4.07–11.33%; coefficients of variation calculated according to Chard 1995). Androgen recovery rates were determined for each sample using tritiated testosterone, with a mean across assays of 64.12%. The minimum detectable androgen concentration was 228.09 pg/mL based on a 16-μL plasma sample cutoff and average recovery rate. Detectable samples had androgen titres ranging from 84.78 to 6025.95 pg/ml. We back-calculated plasma androgen titres from undetectable samples using our assay's minimal detectable levels of 1.95 pg/tube, which yielded a range of 58.46–389.17 pg/mL, depending on plasma volume.

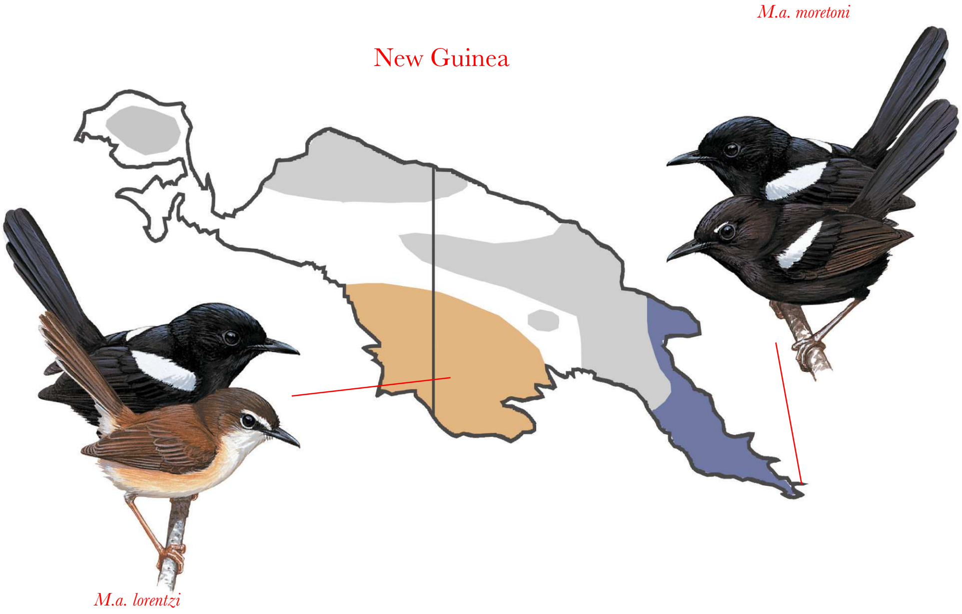

Figure 1
Illustrations of the 2 taxa included in this study (del Hoyo et al. 2017, reproduced with permission). Females are either lacking in melanized ornamentation (**"**unornamented") and the white shoulder patch (left: *M. a. lorentzi*) or possess a melanized black-and-white plumage (**"**ornamented," right: *M. a. moretoni*). Males, illustrated behind females, have similar plumage features in both populations. Gray regions on the map of the island of New Guinea (with the political border drawn in the middle) refer to other populations of White-shouldered Fairywren not included in this study and ranges are based on Birdlife International and NatureServe (2013).

Experimental design

We designed a presentation experiment to test the response of both sexes to the ornamented and unornamented phenotype in each subspecies between January and March, 2016. During this time in both populations, a small number of individuals were nesting or nested after presentation experiments were completed. In each trial, a pair that was known to be territorial was presented with artificial mounts created from cardstock and accompanied by playback consisting of a previously recorded pair duet from their own population. Twelve cardstock mounts were adapted from designs for bird models by an artist (http://www.johanscherft.com). Each mount was painted using a combination of spray paint and gouache paint to resemble the 2 female phenotypes (*n* = 4 ornamented and unornamented each) and male phenotype (*n* = 4) and tail feathers from the respective phenotype were attached to the mount. The use of artificially colored mounts in this study (and others, e.g. Greig et al. 2015) could influence the degree to which aggression measurements reflect interactions between live birds. However, we are interested in the relative difference in aggression between receivers with different phenotypes and interpret our findings in light of comparative aggression differences between females with ornamented or unornamented plumage. We quantified responses to artificial mounts using a male mount with an ornamented female mount or a male mount with an unornamented female mount. We used both a male and a female mount to simulate a territorial intrusion by a White-shouldered Fairywren pair and to distinguish between responses to an ornamented female or a male. Preliminary trials suggested that when presented with a lone female combined with female song, males would often respond alone in an apparent attempt to incite courtship (Enbody ED, unpublished data), so presenting a pair with duet song was used to present a biologically meaningful intrusion in an agonistic context.

Five duets from each population were recorded from pairs using a Marantz PMD 661 Mk II (96 kHz 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). Each duet stimulus for playback consisted of the same duet repeated separated by 10 s, filtered for noise below 500 Hz, and amplitude standardized using Audacity. Playbacks were broadcasted using an iPod touch (Apple Inc, Cupertino, CA) and an Altec Lansing iM227 Orbit MP3 speaker (Altec Lansing LLC, New York, NY).

Prior to running a trial, we located a territorial pair and observed their behavior for 3–5 min prior to placing mounts on a location known to be within their territory. We only tested pairs where both individuals were in adult plumage. We randomly selected a stimulus mount and a local population stimulus song using the following protocol prior to each trial: 1) randomly selected stimulus phenotype (ornamented or unornamented), 2) randomly selected female mount (1–4; of the selected phenotype), 3) randomly selected male mount (1–4), and 4) randomly selected song stimulus (1–5; recorded locally). Each duet stimulus was used an average of 8 times (range: 3–15), each female mount an average of 9 times (range: 6–12), and each male mount an average of 19 times (range: 15–21). We only used songs recorded from pairs not adjacent to the focal territory to avoid the possibility that pairs were familiar with the song stimulus. In other words, each focal pair was presented with either of 2 combinations: ornamented female mount and male mount (with local song) or unornamented female mount and male mount (with local song). This experimental design was chosen to test the response to different visual stimuli, but not to different vocal stimuli.

Male and female mounts were placed on 2 similar 1.5-m tall sticks that we used for all experiments. Each stick included 3 stems 25 cm from the top where one stem held the mount and the 2 other were available for the responding pair to perch on. Male and female mounts were separated by 1 m and a speaker, controlled remotely by observers, was placed in vegetation between and below the mounts. Next, we set up a small blind 20 m away from the mount location for 2 observers. At the start of the trial, 1 observer began recording using the Marantz PMD 661 MK II sound recorder and another using a Sony DCRSX40/L camcorder. Each trial began with 1 min of acclimation time, followed by playback that continued until the focal female came within 1 m of the mount or 5 min had passed, whichever was first. In the former scenario (female approaches within 1 m), 3 more songs were played before ceasing playback, and for the latter scenario playback ceased at the end of 5 min. If the focal female never approached within 10 m after 5 min then the trial was scored as “no response,” ceased, and repeated another day. This usually occurred if the pair had traveled out of hearing distance or was engaged in another territorial dispute. A successful trial ended 10 min after the cessation of playback. Depending on the latency to response to 1 m, trials varied in length between 10 and 15 min (13.34 ± 1.57 min), so continuous time variables are analyzed as rates (duration or # of behavior / trial length). Trials where focal pairs interacted with neighboring territorial holders were discarded (*n* = 9). Following Greig et al. (2015), this protocol allowed us to assess the response to the mounts after a constant number of songs had been played with the pair in close proximity. One observer focused on the female and the other on the male and narrated behaviors into the microphone or video camera, respectively. We only include trials in which the female responded, excluding male-only responses. For both sexes, we recorded the duration of time at 0–5 m from the mounts, the number of duets, flybys (when an individual flew within 0.5 m of the mount), and leapfrogs (when one responding individual hopped over the other responding member of the pair), and latency to each behavior (see Table 1). These behaviors were selected based on their use in previous experimental playback trials for recording aggressive behaviors (Pearson and Rohwer 2000; Uy et al. 2009).

We conducted 75 total presentations (44 in the ornamented population, 31 in the unornamented population), but we only analyzed a sub-set of 55 trials (73%; 33 in the ornamented population, 22 in the unornamented population) in which females responded. Males responded together with the female in 51 of the 55 trials (93% of responsive trials). Raw data on male response in *n* = 15 trials was irretrievably lost in the field due to failure of a data storage device (corresponding data on female response from these trials was not lost). We analyzed our data with and without the 15 trials and obtained qualitatively similar results; for this reason, we decided to include the 15 trails in which only data on female response was available in the analyses presented below.

Statistical analysis

Circulating testosterone

Although all individuals were potentially in breeding condition (see above) we excluded individuals known to be nest building (*n* = 8), egg laying (*n* = 4), incubating (*n* = 11), or nestling provisioning (*n* = 11) stages due to well-documented short term impacts on circulating androgen levels associated with these reproductive activities in this and other species (Lindsay et al. 2009; Schwabl et al. 2014). The remaining dataset included individuals (*n* = 235) in a reproductive stage that we consider to be comparable to temperate

Table 1

The effect of the fixed predictor variables on circulating testosterone using a linear-mixed model, with individual as a random effect

	Value	Std Error	Df	*t*-value	*P*
(Intercept)	4.052	0.386	160	10.509	<0.001*
Subspecies	0.79	0.229	160	3.449	<0.01*
Sex	1.628	0.188	160	8.639	<0.001*
Year (2015)	0.101	0.203	66	0.495	0.622
Year (2016)	1.144	0.309	66	3.698	<0.001*
Julian Date	0.003	0.002	66	1.405	0.165
Subspecies*Sex	−0.98	0.233	160	−4.208	<0.001*
Subspecies*Year(2015)	−0.434	0.287	66	−1.515	0.135
Subspecies*Year(2016)	−0.545	0.332	66	−1.642	0.105

species' "pre-breeding" levels, allowing us to avoid potentially confounding breeding stage effects on testosterone titres in males and females (Lindsay et al. 2009; Schwabl et al. 2014). Age may also influence testosterone levels (Lantz et al. 2017), but our sample sizes for known age birds in the unornamented population were insufficient to include as a predictor in the analysis. Delay between capture and blood sampling has a significant negative effect on circulating testosterone levels in other species (Lindsay et al. 2009), so as noted, we included only samples obtained within 20 min of capture and include time delay as a fixed effect (Lindsay et al. 2009; Lantz et al. 2017). Delay had no significant effect on testosterone titres, so this term was removed from subsequent models (linear regression: $F = 0.023$, df = 233, $R^2 = 0.0001$, $P = 0.878$). Furthermore, the use of playback did not appear to have an effect on circulating testosterone titres (students *t*-test; $t = 1.0459$, df = 127.59, $P = 0.298$). We included individual as a random effect in all models because some individuals were re-captured within and/or between years. After these filtering steps, we analyzed circulating testosterone from a total of 30 adult females and 26 adult males in the population with unornamented females and 84 adult females and 95 adult males in the population with ornamented females.

We first built a generalized linear model with a binomial distribution to test if subspecies differed in the probability that the androgen assay detected testosterone using androgen detection as a binary response variable. We then used linear-mixed models to test whether testosterone levels were related to subspecies, sex, year, time delay (until bled), and Julian date and 2 interaction terms as predictors of testosterone levels (see Supplementary Table S1). Following initial model comparisons, we corrected for heteroscedasticity in the subspecies and sex interaction term using a correction for variance structure. As we were primarily interested in the effect of subspecies and sex on circulating testosterone, we performed stepwise model selection using AIC to remove nonsignificant effects. We present measures of support for our top models based on AIC. The linear-mixed model for testing predictors of testosterone were conducted in R v3.4.2 (R Core Development Team 2017). using *nlme* (Pinheiro et al. 2017).

Response to simulated territorial intrusion

We first tested the effect of song stimulus (i.e. 1–5 at each site) and subspecies on the probability that an individual would respond using the expanded dataset that included both responding individuals and nonresponders. We report the results of a generalized linear model using a binomial distribution on a binary response variable (i.e. "yes" or "no"). We then used principal components analysis (PCA) to quantify responses of free-flying birds from 2 subspecies of White-shouldered Fairywren in our field-based mount presentation experiments. All response variables were log transformed to improve normality, scaled, and centered prior to running the PCA (following Filardi and Smith 2008; Uy et al. 2009). We assessed the effect of subspecies, sex, and stimulus type (and all interactions) on the top 3 principal components using linear-mixed models in the package *lme4* in R (links to R scripts are available from the journal office), because *lme4* handles multiple random effects (Bates et al. 2015). We reduced the effect of pseudoreplication in our results (Kroodsma et al. 2001) by including male mount stimulus (1–4), female mount stimulus (1–4, nested in female phenotype), and song stimulus (1–5, nested in subspecies) as random effects in the model. We removed trials with individuals known to be nest building, incubating, or with nestlings (unornamented $n = 7$, ornamented $n = 5$), because breeding stage can influence aggression in other species (e.g. Hunt et al. 1995). Residuals of the full model were normally distributed and did not violate homoscedastic assumptions. We tested for significant predictors of each model (i.e. the top 3 PCs) using a Wald chi-square test with alpha set at 0.05 using the Anova (Type II) command in the R package "car" (Fox and Weisberg 2011).

RESULTS

Circulating testosterone

A greater proportion of ornamented female samples had detectable levels of testosterone ($n = 42$, 50%) than did unornamented female samples ($n = 9$, 30%), but this effect was not significant (glm: $z = 1.865$, df = 112, $P = 0.062$). The opposite, nonsignificant trend, was found in males, as a lower proportion of males from the population with ornamented females had of detectable testosterone ($n = 76$, 80%) than males from the unornamented population ($n = 25$, 96%): 0.96; glm: $z = -1.743$, df = 119, $P = 0.081$).

The top model for the effect of different predictors on circulating testosterone levels included the interaction between subspecies and sex, the interaction between subspecies and year, and Julian date (Table 1; we report AIC values for stepwise model comparisons in Supplementary Table S1). However, circulating testosterone was only significantly predicted by the interaction between subspecies and sex, and also year; other variables had nonsignificant effects on testosterone (Table 1). Between sex and between-population comparisons (following a Tukey's adjustment for multiple comparisons) of circulating testosterone levels indicate that ornamented females had higher levels of testosterone than unornamented females did, and females of both populations were lower in circulating testosterone than males (Figure 2; Table 2). In contrast, males did not differ between populations in circulating testosterone (Figure 2; Table 2).

Response to simulated territorial intrusion

Stimulus song type (i.e. different song types from the same population that was being presented to; glm females: $z = 0.135$, $P = 0.893$; glm males: $z = 1.719$, $P = 0.086$) did not predict the likelihood that an individual would respond to the simulated territorial intrusion. Neither subspecies was more likely to respond to a trial (glm females: $z = 0.156$, $P = 0.876$; glm males: $z = -0.101$, $P = 0.919$).

Among the 55 trials we included in our analyses, the first 3 PCs cumulatively explained 78.7% of variation in behavioral responses in both sexes to simulated territorial intrusion (eigenvectors and variable loadings are shown in Table 3). We interpret higher values of PC1 as an index of increased aggression, characterized by faster response time and more time spent close to the mount, as well as greater rates of pair coordination behavior (duets and leapfrogs). Interpretation is less clear for PC2, but the component seems to correspond to responses that were rapid, but of short duration, or characterize the overall motivation of an individual to respond. PC3 was also associated with increased aggressive response, in particular a high number of flybys, and by less pair coordination behavior (duets and leapfrogs). Individual response variables are shown in Supplementary Figure S2.

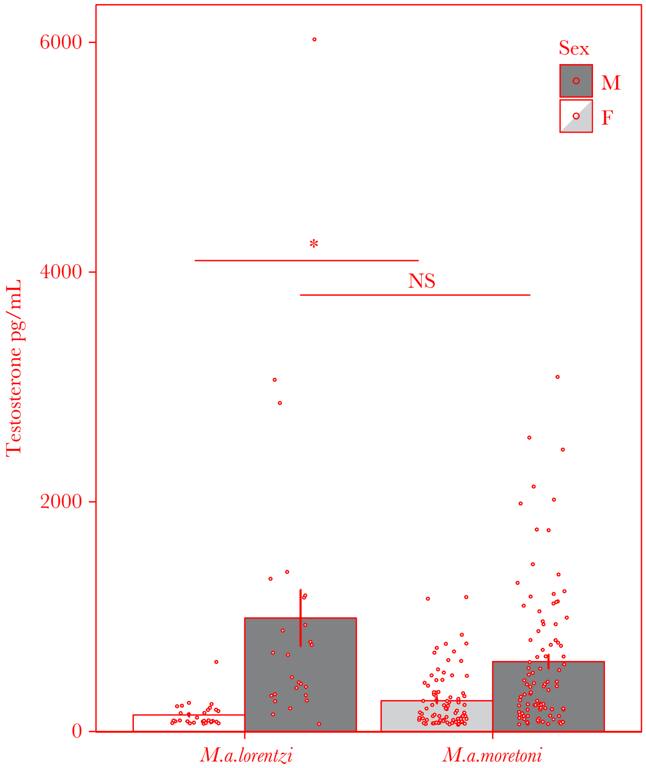

Figure 2
The relationship between circulating testosterone titres (untransformed) and subspecies in male (gray) and female (white or light gray) White-shouldered Fairywrens. Error bars denote standard error and significant comparisons are denoted by an asterisk (*$P < 0.01$).

Subspecies was a significant predictor of PC1, PC2, and PC3 (Table 4; Figure 3), indicating that birds in the population with ornamented females exhibited a more aggressive response to simulated territorial intrusion than birds in the population with unornamented females did. Responses were tightly correlated between the sexes, consistent with the observation that pairs generally responded together with similar intensity. However, sex was a significant predictor of PC2, indicating that, between the 2 populations, females may have been overall less quick to approach and remain close to the mount, while still interacting to a high degree in pair coordination behavior and song. No other effects, including stimulus type or any interactions, were significant predictors of PC1, PC2, or PC3 (Table 4).

DISCUSSION

This study examined the association between circulating testosterone, ornamentation, and aggression in females and males of 2 recently diverged populations of passerine bird, White-shouldered Fairywrens (*Malurus alboscapulatus*), that differ in female ornamentation. In doing so, our goal was to explore the extent to which the integrated ornamented phenotype is conserved between the sexes and across populations with variable ornamentation. We found that females with an ornamented plumage phenotype have higher levels of circulating testosterone than do females with an unornamented phenotype. Females and males from the population with elevated levels of female testosterone and female ornamentation also exhibited greater aggressive response to experimental territorial intrusions. Our results indicate that a hormonal mechanism mediates integration of plumage ornamentation and aggression in female White-shouldered Fairywrens, following the same associations between these traits previously documented in male Red-backed Fairywrens (Lindsay et al. 2009; Webster et al. 2010; Lindsay et al. 2011, but see Barron et al. 2015). Importantly, we recovered no corresponding relationship between testosterone and aggression among male White-shouldered Fairywrens suggesting that selection acts on female traits independent of selection on males.

Our finding that 2 populations with variation in female ornamentation differ in circulating levels of testosterone suggests a role for testosterone in mediating the female ornamented phenotype. In other sexually dimorphic species, testosterone-implanted females produce only a portion of the male's morphological phenotype

Table 2
Contrasts for linear-mixed model comparing circulating testosterone between subspecies and sex following a Tukey's adjustment for multiple comparisons

Sex	Subspecies	estimate	SE	Df	t.ratio	*P*
F–F	*M. a. lorentzi*–*M. a. moretoni*	−0.464	0.136	66	−3.4	<0.01*
M–F	*M. a. lorentzi*–*M. a. lorentzi*	−1.628	0.188	160	−8.639	<0.0001*
	M. a. lorentzi–*M. a. moretoni*	−1.111	0.148	66	−7.49	<0.0001*
	M. a. moretoni–*M. a. lorentzi*	−1.164	0.210	66	−5.542	<0.0001*
	M. a. moretoni–*M. a. moretoni*	−0.648	0.137	160	−4.731	<0.0001*
M–M	*M. a. lorentzi*–*M. a. moretoni*	0.516	0.218	66	2.368	0.0935

M. a. moretoni have ornamented females and *M. a. lorentzi* have unornamented females.

(*Malurus cyaneus*, Peters 2007; *Malurus melanocephalus*, Lindsay et al. 2016) or song type (*Sturnis vulgaris*, Ridder et al. 2002). Together, these results are consistent with circulating testosterone being partially responsible for regulating sexual dimorphism, suggesting that other circulating hormones (e.g. oestrogen, Owens and Short 1995) or hormone receptor differences (Rosvall et al. 2012) also mediate ornament expression differences between the sexes. Yet in White-shouldered Fairywrens, the association between elevated plasma testosterone, the production of an ornamented plumage, and increased aggression implies similarity in the relationship between testosterone and ornament production between the sexes. This finding supports the supposition that differential sex steroid production and secretion in adulthood is a potential mechanism through which sexual dimorphism can be maintained, despite males and females sharing most of their genetic architecture (Adkins-Regan 2005). Despite our robust findings, the pattern we observe is limited to 2 allopatric populations and is therefore limited by overall sample size. Thus, though the differences detected in our statistical analysis appear sufficient for documenting the general patterns between the different female phenotypes, the complete pathway for ornament production is still unknown, but could be further addressed using experimental testosterone-implant studies.

Table 3

PCA eigenvectors (standard deviation), variance explained, and loadings for the top 3 PCs

	PC1	PC2	PC3
Standard deviation	1.462	1.283	0.968
Proportion of Variance	0.356	0.275	0.156
Leapfrogs	0.377	0.363	−0.458
Duets	0.47	0.456	0.271
Time <5 m	0.383	−0.441	−0.442
Latency to 5 m	−0.367	0.524	0.149
Latency to first duet	−0.566	−0.154	−0.257
Flybys	0.188	−0.41	0.659

Leapfrogs is the rate of leapfrog behavior in each trial, duets are the rate of coordinated songs per trial, time <5 m is the proportion of time spent at a distance of <5 m from the mounts, and flybys are flights within 0.5 m of the mounts.

Intriguingly, the observation that testosterone is elevated in ornamented females, but not males of that population, suggests that testosterone is not elevated in females simply as a byproduct of selection for higher androgen levels in males (Ketterson et al. 2005; Møller et al. 2005). Accordingly, adaptive explanations for the appearance of ornamentation and higher aggression should be considered for the evolution of ornamentation in females. Female aggression can have fitness consequences in both vertebrates (Dloniak et al. 2006; Stockley and Bro-Jørgensen 2011; Cain and Langmore 2016) and invertebrates (Elias et al. 2010; Bath et al. 2017). Female Red-backed Fairywrens also respond aggressively to female intruders (Karubian et al. 2008). Therefore, it is parsimonious to suggest that female ornaments would function in similar contexts where they are present in White-shouldered Fairywrens. Aggression is important for establishing dominance relationships and mediating intrasexual conflict, a process that is associated with sexual selection and speciation (Lipshutz 2018). Our study demonstrates that aggressive behavior differs between females of 2 populations that vary in female, but not male ornamentation. We suggest that female ornaments in this system function in territory defense and may mediate access to limiting resources (ecological or social) by gaining access to higher quality territories. In addition, heavily weighted principal components for duets and leapfrog behavior suggest that ornamented females may be coordinating to a greater degree with their mate to improve territory defense. This is important, because both males and females in the population with ornamented females responded with greater aggression scores, so pair coordination (and not just female aggression alone) is likely a strong driver of the observed population differences in aggression. If this hypothesis is true, social selection may play a role in shaping female phenotypic traits by improving access to ecological or

Table 4

Model estimates, standard error, and *t*-value for the effect of subspecies (focal *M. a. moretoni* or focal *M. a. lorentzi*), sex, mount stimulus (ornamented or unornamented), and all interactions on the top 3 principal components

		Estimate	Standard error	*t*-value	Chi square	df	*P*
PC1	Subspecies	1.004	0.575	1.745	27.002	1	<0.01*
	Sex	−0.498	0.664	−0.75	0.129	1	0.72
	Stimulus	−0.688	0.601	−1.144	0.886	1	0.346
	Subspecies*Sex	0.789	0.85	0.928	0.058	1	0.81
	Subspecies*Stimulus	0.774	0.754	1.027	0.209	1	0.648
	Sex*Stimulus	0.627	0.944	0.664	0.087	1	0.768
	Subspecies*Sex*Stimulus	−1.26	1.187	−1.061	1.127	1	0.289
PC2	Subspecies	0.635	0.566	1.12	4.846	1	<0.05*
	Sex	−0.765	0.564	−1.356	6.445	1	<0.05*
	Stimulus	−0.281	0.591	−0.476	0.364	1	0.546
	Subspecies*Sex	0.341	0.726	0.47	0.06	1	0.806
	Subspecies*Stimulus	0.175	0.677	0.258	0	1	0.988
	Sex*Stimulus	0.145	0.811	0.179	0.065	1	0.799
	Subspecies*Sex*Stimulus	−0.426	1.021	−0.417	0.174	1	0.676
PC3	Subspecies	0.872	0.379	2.298	19.151	1	<0.01*
	Sex	−0.253	0.388	−0.653	2.365	1	0.124
	Stimulus	0.139	0.445	0.313	0.173	1	0.678
	Subspecies*Sex	0.27	0.499	0.541	0.044	1	0.834
	Subspecies*Stimulus	0.185	0.465	0.397	0.007	1	0.931
	Sex*Stimulus	−0.078	0.558	−0.14	0.921	1	0.337
	Subspecies*Sex*Stimulus	−0.389	0.703	−0.553	0.306	1	0.58

Significance tested using a Wald's chi-square test and we report chi-square values, degrees of freedom, and *P* value with significant effects marked with an asterisk. Subspecies is a significant predictor of PC1, PC2, and PC3.

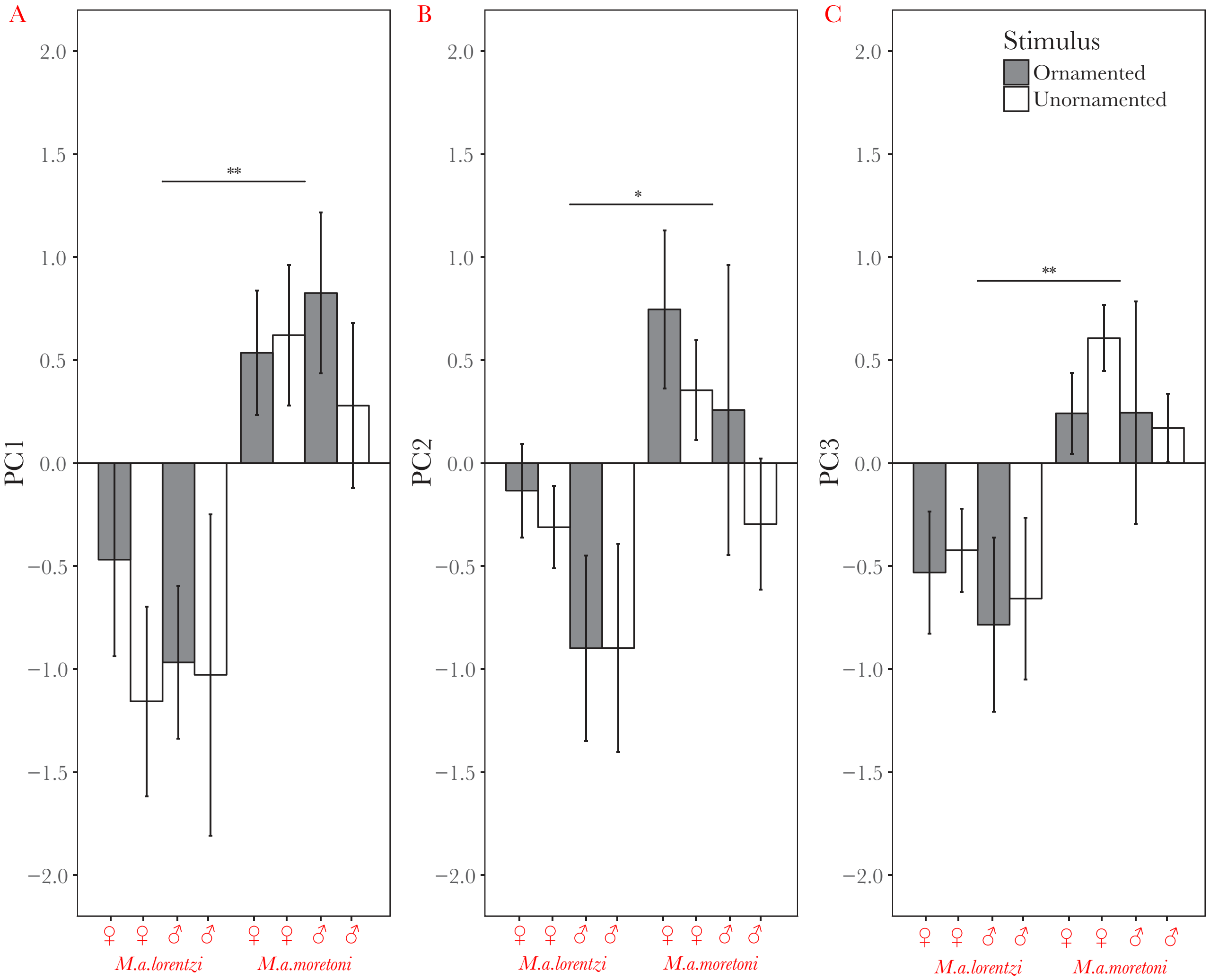

Figure 3

Response of White-shouldered Fairywrens to simulated territorial intrusions. Shown are responses by free-flying females and males of 2 populations that differ in degree of female ornamentation. In *M. a. lorentzi*, females are unornamented and, in *M. a. moretoni*, females are ornamented. Color of the bar represents the type of stimulus presented: white denotes trials where an unornamented stimulus was presented and gray denotes trials where an ornamented stimulus was presented. Both sexes are plotted on the *x* axis. Individuals of both sexes in the population with ornamented females responded with greater values in PC1 (A), PC2 (B), and PC3 (C). Females responded with greater overall values of PC2 than males. Error bars show standard error and significance is shown for comparisons between populations (**$P < 0.01$, *$P < 0.05$).

social resources (West-Eberhard 1979). It is possible that habitat or environmental differences drive this pattern, as the site with unornamented females shows greater variability in seasonality of rainfall (see Methods). For example, resources (e.g. food, nesting sites, or nest and display materials) may be more evenly distributed in the less seasonal environment where the ornamented population occurs and species that experience little seasonality may have low territory turnover between years (Stutchbury and Morton 2001), which could lead to long-term investment in territory maintenance and elevate the importance of territory defense behaviors.

Our findings provide compelling evidence that males and female in a population with ornamented females respond with a greater degree of territorial defense than individuals in a population with unornamented females. Differential aggression towards heterospecific stimuli can influence species interactions by mediating aggressive interactions (e.g. Jankowski et al. 2010; Tobias et al. 2014; Freeman 2016), but we find that presentation of different stimulus types (ornamented or unornamented female) does not elicit differential aggression in either population. The artificial mounts used in this project may elicit a different amplitude of response compared with live birds, but the lack of differential response to different stimuli at least suggests that broader patterns in aggression represents a true difference in baseline aggression in the 2 subspecies, rather than a differential response to the stimulus type we presented. These results also suggest that unornamented females do not have a pre-existing bias to respond with greater aggression towards a perceived aggressive phenotype (e.g. through sensory drive: Endler 1992). However, future research investigating responses to song types versus plumage phenotypes from the 2 populations would clarify whether there are pre-existing biases for signal types in White-shouldered Fairywrens.

Male testosterone levels were not elevated in the White-shouldered Fairywren population that showed higher male and female territorial responses, and in which females had higher testosterone levels. Our findings suggest that in White-shouldered Fairywrens, as in Red-backed Fairywrens (Karubian et al. 2011), testosterone levels are elevated in males relative to females, but that relative subspecific differences in testosterone for males do not mirror differences in aggression. Similar research in songbirds with differing aggression between populations has suggested that testosterone sensitivity may have a more direct effect than differences in circulating levels (Bergeon Burns et al. 2013; Bergeon Burns et al. 2014), which may apply to male White-shouldered Fairywrens as well. It remains to be shown whether sex differences in White-shouldered Fairywrens suggest decoupling of male territorial behavior from testosterone levels or result from differences in habitat structure, breeding density, and seasonality that could affect male testosterone levels and territorial behavior. For example, males in the population with unornamented females appear to spend more time interacting with members of other groups than in the population with ornamented females (Enbody ED, J. Boersma J, unpublished data), which may elevate testosterone levels (e.g. Challenge Hypothesis; Wingfield et al. 1990), though this potential relationship requires additional work.

CONCLUSIONS

The multiple effects of testosterone can couple different components of an individual's physiology, morphology, and behavior to produce an integrated phenotypes (Ketterson et al. 2009). Findings in this study corroborate prior work on fairywren males suggesting that testosterone, ornamentation, and behavior are interrelated in the production of competitive reproductive phenotypes (Wingfield et al. 1990; Hau 2001; Archer 2006). The gain or loss of ornamentation in females, independent of males, has been shown to influence macroevolutionary speciation patterns (Irwin 1994; Burns 1998; Figuerola and Green 2000; Friedman et al. 2009; Price and Whalen 2009; Johnson et al. 2013; Shultz and Burns 2017) and our study suggests that it similarly influences differentiation among populations. Evidence of associations between population-specific female ornamentation, testosterone levels, and aggressive behavior suggests that females hormonally integrate morphological and competitive traits using a similar process as males in other species (namely testosterone, Cain and Ketterson 2012). Importantly, transitions in the female integrated phenotype in White-shouldered Fairywrens occur independently of males and are reflected in behavioral differences between populations, which is consistent with selection on female ornamentation involved in competition over territorial resources. This lays out promising avenues for further study of female ornament evolution focusing on mechanisms and underlying components of integrated male and female phenotypes.

SUPPLEMENTARY MATERIAL

Supplementary data are available at *Behavioral Ecology* online.

FUNDING

This work was supported by the National Science Foundation (IOS-1354133 to J.K.); the Disney Worldwide Conservation Fund (to J.K.); the American Ornithologist's Union (to E.D.E.); the American Museum of Natural History Chapman Memorial Fund (to E.D.E.); the National Geographic Society Young Explorers Grant (to E.D.E.); the Animal Behavior Society (to E.D.E.); the Tulane University Ecology & Evolutionary Biology Department Student Research Grant and the Gunning Fund (to E.D.E.); and the Washington State University Elling Research Grant (to J.P.B.).

Field data collection was made possible by P. Chaon, S. Ketaloya, D. Nason, M. Saiga, K. Saiga, G. Kareba, R. Biggoneau, M. Aliaga, B. Barkley, and B. Magnier. H. Rust assisted with audio transcription and G. Sprehn assisted with experimental mounts construction. The manuscript was improved by feedback from M. J. Blum, E. P. Derryberry, M. S. Webster, M. J. West-Eberhard, B. Geary, S. Lantz, J. Jones, S. Khalil, Z. Diaz-Martin, L. Browne, and K. Narasimhan. Papua New Guinea research permits were obtained through the assistance of the National Research Institute. We also thank the Milne Bay and Western Province provincial governments, and local level governments and land owners at each site for permits and permissions for field data collection.

Data accessibility: Analyses reported in this article can be reproduced using the data provided by Enbody et al. (2018).

Handling editor: Naomi Langmore

REFERENCES

Adkins-Regan E. 2005. Hormones and animal social behavior. Princeton (NJ): Princeton University Press.

Amundsen T. 2000. Why are female birds ornamented? Trends Ecol Evol. 15:149–155.

Amundsen T, Parn H. 2006. Female coloration: review of functional and nonfunctional hypotheses. In: Hill GE, Mcgraw KJ, editors. Bird coloration, Vol. II: function and evolution. Cambridge (MA): Harvard University Press. p. 280–345.

Archer J. 2006. Testosterone and human aggression: an evaluation of the challenge hypothesis. Neurosci Biobehav Rev. 30:319–345.

Barron DG, Webster MS, Schwabl H. 2015. Do androgens link morphology and behaviour to produce phenotype-specific behavioural strategies? Anim Behav. 100:116–124.

Bates D, Mächler M, Bolker B, Walker S. 2015. Fitting linear mixed-effects models using lme4. J Stat Softw. 67:1–48.

Bath E, Bowden S, Peters C, Reddy A, Tobias JA, Easton-Calabria E, Seddon N, Goodwin SF, Wigby S. 2017. Sperm and sex peptide stimulate aggression in female Drosophila. Nat Ecol Evol. 1:0154.

Bergeon Burns CM, Rosvall KA, Hahn TP, Demas GE, Ketterson ED. 2014. Examining sources of variation in HPG axis function among individuals and populations of the dark-eyed junco. Horm Behav. 65:179–187.

Bergeon Burns CM, Rosvall KA, Ketterson ED. 2013. Neural steroid sensitivity and aggression: comparing individuals of two songbird subspecies. J Evol Biol. 26:820–831.

Birdlife International and NatureServe. 2013. Bird species distribution maps of the world. Available from: http://www.birdlife.org (Accessed 4 April 2014).

Buchanan KL, Cockburn A. 2013. Fairy-wrens and their relatives (Maluridae) as model organisms in evolutionary ecology: the scientific legacy of ian Rowley and Eleanor Russell. Emu. 113:i–vii.

Burns KJ. 1998. A phylogenetic perspective on the evolution of sexual dichromatism in tanagers (thraupidae): the role of female versus male plumage. Evolution. 52:1219–1224.

Cain KE, Ketterson ED. 2012. Competitive females are successful females; phenotype, mechanism, and selection in a common songbird. Behav Ecol Sociobiol. 66:241–252.

Cain KE, Langmore NE. 2016. Female song and aggression show contrasting relationships to reproductive success when habitat quality differs. Behav Ecol Sociobiol. 70:1–11.

Clutton-Brock TH, Huchard E. 2013. Social competition and selection in males and females. Philos Trans R Soc Lond B Biol Sci. 368:20130074.

Cox RM, Cox CL, McGlothlin JW, Card DC, Andrew AL, Castoe TA. 2017. Hormonally mediated increases in sex-biased gene expression accompany the breakdown of between-sex genetic correlations in a sexually dimorphic lizard. Am Nat. 189:315–332.

Cox CL, Hanninen AF, Reedy AM, Cox RM. 2015. Female anoles retain responsiveness to testosterone despite the evolution of androgen-mediated sexual dimorphism. Funct Ecol. 29:758–767.

Cox RM, Zilberman V, John-Alder HB. 2008. Testosterone stimulates the expression of a social color signal in Yarrow's Spiny lizard, *Sceloporus jarrovii*. J Exp Zool A Ecol Genet Physiol. 309:505–514.

Crowhurst CJ, Zanollo V, Griggio M, Robertson J, Kleindorfer S. 2012. White flank spots signal feeding dominance in female diamond firetails, *Stagonopleura guttata*. Ethology. 118:63–75.

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

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

Dijkstra PD, Verzijden MN, Groothuis TG, Hofmann HA. 2012. Divergent hormonal responses to social competition in closely related species of haplochromine cichlid fish. Horm Behav. 61:518–526.

Dloniak SM, French JA, Holekamp KE. 2006. Rank-related maternal effects of androgens on behaviour in wild spotted hyaenas. Nature. 440:1190–1193.

Dowling J, Webster MS. 2017. Working with what you've got: unattractive males show greater mate-guarding effort in a duetting songbird. Biol Lett. 13:20160682.

Driskell AC, Norman JA, Pruett-Jones S, Mangall E, Sonsthagen S, Christidis L. 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.

Elias DO, Botero CA, Andrade MCB, Mason AC, Kasumovic MM. 2010. High resource valuation fuels "desperado" fighting tactics in female jumping spiders. Behav Ecol. 21:868–875.

Emlen DJ, Marangelo J, Ball B, Cunningham CW. 2005. Diversity in the weapons of sexual selection: horn evolution in the beetle genus *Onthophagus* (Coleoptera: Scarabaeidae). Evolution. 59:1060–1084.

Enbody ED, Boersma J, Schwabl H, Karubian J. 2018. Data from: female ornamentation is associated with elevated aggression and testosterone in a tropical songbird. Dryad Digital Repository. http://dx.doi.org/10.5061/dryad.qn037tk.

Enbody ED, Lantz SM, Karubian J. 2017. Production of plumage ornaments among males and females of two closely related tropical passerine bird species. Ecol Evol. 7:4024–4034.

Endler JA. 1992. Sensory drive. Does sensory drive biology bias or constrain the direction of evolution? Am Nat. 139:153.

Fernald RD. 1976. The effect of testosterone on the behavior and coloration of adult male cichlid fish (*Haplochromis burtoni*, Günther). Horm Res. 7:172–178.

Figuerola J, Green AJ. 2000. The evolution of sexual dimorphism in relation to mating patterns, cavity nesting, insularity and sympatry in the Anseriformes. Funct Ecol. 14:701–710.

Filardi CE, Smith CE. 2008. Social selection and geographic variation in two monarch flycatchers from the solomon islands. Condor. 110:24–34.

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

Freeman BG. 2016. Strong asymmetric interspecific aggression between two sympatric New Guinean robins. Ibis. 158:75–81.

Friedman NR, Hofmann CM, Kondo B, Omland KE. 2009. Correlated evolution of migration and sexual dichromatism in the New World orioles (*icterus*). Evolution. 63:3269–3274.

Goymann W, Schwabl I, Trappschuh M, Hau M. 2007. Use of ethanol for preserving steroid and indoleamine hormones in bird plasma. Gen Comp Endocrinol. 150:191–195.

Goymann W, Wingfield JC. 2014. Male-to-female testosterone ratios, dimorphism, and life history - what does it really tell us? Behav Ecol. 25:685–699.

Greig EI, Baldassarre DT, Webster MS. 2015. Differential rates of phenotypic introgression are associated with male behavioral responses to multiple signals. Evolution. 69:2602–2612.

Hau M. 2001. Timing of breeding in variable environments: tropical birds as model systems. Horm Behav. 40:281–290.

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

Hofmann CM, Cronin TW, Omland KE. 2008. Evolution of sexual dichromatism. 1. Cnvergent losses of elaborate female coloration in new world orioles (*Icterus* spp.). Auk. 125:778–789.

del Hoyo J, Elliott A, Sargatal J, Christie DA, de Juana E, editors. 2017. Handbook of the birds of the world alive. Barcelona: Lynx Edicions.

Hunt K, Wingfield JC, Astheimer LB, Buttemer WA, Hahn TP. 1995. Temporal patterns of territorial behavior and circulating testosterone in the lapland longspur and other arctic passerines. Integr Comp Biol. 35:274–284.

Irwin RE. 1994. The evolution of plumage dichromatism in the New World Blackbirds: social selection on female brightness. Am Nat. 144:890.

Jankowski JE, Robinson SK, Levey DJ. 2010. Squeezed at the top: interspecific aggression may constrain elevational ranges in tropical birds. Ecology. 91:1877–1884.

Jawor JM, Young R, Ketterson ED. 2006. Females competing to reproduce: dominance matters but testosterone may not. Horm Behav. 49:362–368.

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

Karubian J. 2002. Costs and benefits of variable breeding plumage in the red-backed fairy-wren. Evolution. 56:1673–1682.

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, Lindsay WR, Schwabl H, Webster MS. 2011. Bill coloration, a flexible signal in a tropical passerine bird, is regulated by social environment and androgens. Anim Behav. 81:795–800.

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

Ketterson ED, Atwell JW, McGlothlin JW. 2009. Phenotypic integration and independence: hormones, performance, and response to environmental change. Integr Comp Biol. 49:365–379.

Ketterson ED, Nolan VJ. 1992. Hormones and life histories: an integrative approach. Am Nat. 140:S33–S62.

Ketterson ED, Nolan V, Sandell M. 2005. Testosterone in females: mediator of adaptive traits, constraint on sexual dimorphism, or both? Am Nat. 166:S85–S98.

Kimball RT, Ligon JD. 1999. Evolution of avian plumage dichromatism from a proximate perspective. Am Nat. 154:182–193.

Kriner E, Schwabl H. 1991. Control of winter song and territorial aggression of female robins (*Erithacus rubecula*) by testosterone. Ethology. 87:37–44.

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

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

Langmore NE, Cockrem JF, Candy EJ. 2002. Competition for male reproductive investment elevates testosterone levels in female dunnocks, *Prunella modularis*. Proc Biol Sci. 269:2473–2478.

Lank DB, Coupe M, Wynne-Edwards KE. 1999. Testosterone-induced male traits in female ruffs (*Philomachus pugnax*): autosomal inheritance and gender differentiation. Proc R Soc B. 266:2323–2330.

Lantz SM, Boersma J, Schwabl H, Karubian J. 2017. Early-moulting red-backed fairywren males acquire ornamented plumage in the absence of elevated androgens. Emu. 117:170–180.

Lee JY, Joseph L, Edwards SV. 2012. A species tree for the Australo-Papuan Fairy-wrens and allies (Aves: Maluridae). Syst Biol. 61:253–271.

Lindsay WR, Barron DG, Webster MS, Schwabl H. 2016. Testosterone activates sexual dimorphism including male-typical carotenoid but not melanin plumage pigmentation in a female bird. J Exp Biol. 219:3091–3099.

Lindsay WR, Webster MS, Schwabl H. 2011. Sexually selected male plumage color is testosterone dependent in a tropical passerine bird, the red-backed fairy-wren (*Malurus melanocephalus*). PLoS One. 6:e26067.

Lindsay WR, Webster MS, Varian CW, Schwabl H. 2009. Plumage colour acquisition and behaviour are associated with androgens in a phenotypically plastic tropical bird. Anim Behav. 77:1525–1532.

Lipshutz SE. 2018. Interspecific competition, hybridization, and reproductive isolation in secondary contact: missing perspectives on males and females. Curr Zool. 64:75–88.

Møller AP, Garamszegi LZ, Gil D, Hurtrez-Boussès S, Eens M. 2005. Correlated evolution of male and female testosterone profiles in birds and its consequences. Behav Ecol Sociobiol. 58:534–544.

Mougeot F, Irvine JR, Seivwright L, Redpath SM, Piertney S. 2004. Testosterone, immunocompetence, and honest sexual signaling in male red grouse. Behav Ecol. 15:930–937.

Muck C, Goymann W. 2011. Throat patch size and darkness covaries with testosterone in females of a sex-role reversed species. Behav Ecol. 22:1312–1319.

Van Oordt GJ, Junge GCA. 1934. The relation between the gonads and the secondary sexual characters in the ruff (*Philomachus pugnax*). Latv. biol. Biedr. Raksti:141–145.

Owens IP, Short RV. 1995. Hormonal basis of sexual dimorphism in birds: implications for new theories of sexual selection. Trends Ecol Evol. 10:44–47.

Pärn H, Lindström KM, Sandell M, Amundsen T. 2008. Female aggressive response and hormonal correlates - An intrusion experiment in a free-living passerine. Behav. Ecol Sociobiol. 62:1665–1677.

Pearson S, Rohwer SA. 2000. Asymmetries in male aggression across an avian hybrid zone. Behav Ecol. 11:93–101.

Peters A. 2007. Testosterone treatment of female Superb Fairy-wrens *Malurus cyaneus* induces a male-like prenuptial moult, but no coloured plumage. Ibis. 149:121–127.

Peters A, Astheimer LB, Boland CRJ, Cockburn A. 2000. Testosterone is involved in acquisition and maintenance of sexually selected male plumage in superb fairy-wrens, *Malurus cyaneus*. Behav Ecol Sociobiol. 47:438–445.

Pigliucci M. 2003. Phenotypic integration: studying the ecology and evolution of complex phenotypes. Ecol Lett. 6:265–272.

Pinheiro J, Bates D, DebRoy S, Sarkar D and R Core Team. 2017. _nlme: Linear and Nonlinear Mixed Effects Models_. R package version 3.1-131. https://CRAN.R-project.org/package=nlme

Price JJ, Whalen LM. 2009. Plumage evolution in the oropendolas and caciques: different divergence rates in polygynous and monogamous taxa. Evolution. 63:2985–2998.

Pryke SR. 2007. Fiery red heads: female dominance among head color morphs in the Gouldian finch. Behav Ecol. 18:621–627.

R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Ridder DE, Pinxten R, Eens M. 2002. Short- and long-term effects of male-like concentrations of testosterone on female European starlings (*Sturnus vulgaris*). Auk. 119:487–497.

Roberts ML, Ras E, Peters A. 2009. Testosterone increases UV reflectance of sexually selected crown plumage in male blue tits. Behav Ecol. 20:535–541.

Rosvall KA. 2013. Proximate perspectives on the evolution of female aggression: good for the gander, good for the goose? Philos Trans R Soc Lond B Biol Sci. 368:20130083.

Rosvall KA, Bergeon Burns CM, Barske J, Goodson JL, Schlinger BA, Sengelaub DR, Ketterson ED. 2012. Neural sensitivity to sex steroids predicts individual differences in aggression: implications for behavioural evolution. Proc R Soc B. 279:3547–3555.

Rosvall KA, Bergeon Burns CM, Jayaratna SP, Ketterson ED. 2016. Divergence along the gonadal steroidogenic pathway: implications for hormone-mediated phenotypic evolution. Horm Behav. 84:1–8.

Rowley I, Russell E. 1997. Fairy-wrens and grasswrens: maluridae. Oxford: Oxford University Press.

Rubenstein DR, Lovette IJ. 2009. Reproductive skew and selection on female ornamentation in social species. Nature. 462:786–789.

Sandell MI. 2007. Exogenous testosterone increases female aggression in the European starling (*Sturnus vulgaris*). Behav Ecol Sociobiol. 62:255–262.

Schwabl H, Dowling J, Baldassarre DT, Gahr M, Lindsay WR, Webster MS. 2015. Variation in song system anatomy and androgen levels does not correspond to song characteristics in a tropical songbird. Anim Behav. 104:39–50.

Schwabl H, Lindsay WR, Barron DG, Webster MS. 2014. Endocrine correlates of mate choice and promiscuity in females of a socially monogamous avian mating system with alternative male reproductive phenotypes. Curr Zool. 60:804–815.

Shultz AJ, Burns KJ. 2017. The role of sexual and natural selection in shaping patterns of sexual dichromatism in the largest family of songbirds (Aves: Thraupidae). Evolution. 71:1061–1074.

Simmons LW, Emlen DJ. 2008. No fecundity cost of female secondary sexual trait expression in the horned beetle Onthophagus sagittarius. J Evol Biol. 21:1227–1235.

Staub NL, De Beer M. 1997. The role of androgens in female vertebrates. Gen Comp Endocrinol. 108:1–24.

Stockley P, Bro-Jørgensen J. 2011. Female competition and its evolutionary consequences in mammals. Biol Rev Camb Philos Soc. 86:341–366.

Stutchbury BJM, Morton ES. 2001. Behavioral ecology of tropical birds. Academic Press Cambridge, MA.

Tobias JA, Cornwallis CK, Derryberry EP, Claramunt S, Brumfield RT, Seddon N. 2014. Species coexistence and the dynamics of phenotypic evolution in adaptive radiation. Nature. 506:359–363.

Uy JA, Moyle RG, Filardi CE. 2009. Plumage and song differences mediate species recognition between incipient flycatcher species of the Solomon Islands. Evolution. 63:153–164.

Veiga JP, Polo V. 2008. Fitness consequences of increased testosterone levels in female spotless starlings. Am Nat. 172:42–53.

Webster MS, Karubian J, Schwabl H. 2010. Dealing with uncertainty. Flexible reproductive strategies by a tropical passerine bird in an unstable ecological and social environment. Adv Study Behav. 42:123–153.

Weiss SL, Kennedy EA, Bernhard JA. 2009. Female-specific ornamentation predicts offspring quality in the striped plateau lizard, *Sceloporus virgatus*. Behav Ecol. 20:1063–1071.

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

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

Wingfield JC, Hegner RE, Dufty AM, Ball GF. 1990. The "challenge hypothesis": theoretical implications for patterns of testosterone secretion, mating systems, and breeding strategies. Am Nat. 136:829–846.

Zysling DA, Greives TJ, Breuner CW, Casto JM, Demas GE, Ketterson ED. 2006. Behavioral and physiological responses to experimentally elevated testosterone in female dark-eyed juncos (*Junco hyemalis carolinensis*). Horm Behav. 50:200–207.