Re-examining a standard proxy for sexual selection: plumage dichromatism is negatively correlated with song divergence in passerine birds

Christopher R. Cooney[1,2*], Hannah E. A. MacGregor[1,3], Joseph A. Tobias[1,4] & Nathalie Seddon[1]

[1]*Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK*.
[2]*Department of Animal and Plant Sciences, University of Sheffield, Western Bank, Sheffield S10 2TN, UK.*
[3]*School of Biological Sciences, University of Bristol, Bristol BS8 1TQ, UK.*
[4]*Department of Life Sciences, Imperial College London, Silwood Park, Buckhurst Road, Ascot, Berkshire, SL5 7PY, UK.*

***Author for correspondence:** c.cooney@sheffield.ac.uk

Article type: Research article
Running head: Sexual dichromatism and song divergence in birds
Number of words in abstract: 200
Total number of words in main text: 4166
Number of figures: 3 (10 in Supplementary Information)
Number of tables: 1 (11 in Supplementary Information)
Number of references: 77

Author contributions: C.R.C., J.A.T and N.S developed the conceptual framework. C.R.C, J.A.T and H.E.A.M collected data. C.R.C conducted the analyses. C.R.C wrote the manuscript, with input from all authors.

Abstract

Sexual selection is proposed to be an important driver of speciation and phenotypic diversification in animal systems. However, previous phylogenetic tests have produced conflicting results, perhaps because they have focused on a single signalling modality (visual ornaments), whereas sexual selection may act on alternative signalling modalities (e.g. acoustic ornaments). Here we compile phenotypic data from 259 avian sister species pairs to assess the relationship between visible plumage dichromatism—a standard index of sexual selection in birds—and macroevolutionary divergence in the other major avian signalling modality: song. We find evidence for a strong negative relationship between the degree of plumage dichromatism and divergence in song traits, which remains significant even when accounting for other key factors, including habitat type, ecological divergence and interspecific interactions. This negative relationship is opposite to the pattern expected by a straightforward interpretation of the sexual selection-diversification hypothesis, whereby higher levels of dichromatism indicating strong sexual selection should be related to greater levels of mating signal divergence regardless of signalling modality. Our findings imply a 'trade-off' between the elaboration of visual ornaments and the diversification of acoustic mating signals, and suggest that the effects of sexual selection on diversification can only be determined by considering multiple alternative signalling modalities.

Keywords: birds, dichromatism, divergence, plumage, sexual selection, song, trade-off

Introduction

Previous studies have provided theoretical and empirical evidence that sexual selection can stimulate the rapid divergence of traits involved in mate choice and species recognition [1, 2], supporting the longstanding view that sexual selection is an important driver of speciation and lineage diversification [3-6]. However, direct support for this hypothesis is relatively weak and inconsistent among taxa [7], with a series of studies finding no evidence of significant correlations between sexual selection and either species richness or speciation rate when studied across birds [8-11], mammals, butterflies and spiders [12], and certain fish taxa [13]. Whereas these observations suggest that sexual selection has limited effects on diversification at macroevolutionary scales, an alternative possibility is that standard comparative analyses are simply ineffective because they rely on crude phenotypic proxies to estimate variation in sexual selection across disparate species.

To quantify the intensity of sexual selection, most existing large-scale studies in birds have used visible sex-differences in plumage colouration [e.g. 2, 8, 9, 10, 14, 15]. This metric—usually termed 'plumage dichromatism'—has become a standard proxy for sexual selection because it is easily measured and positively correlated with other indices of sexual selection such as testes size, the degree of polygyny, and the frequency of extra-pair paternity [16-19]. Nonetheless, the extent to which plumage dichromatism provides an accurate and consistent estimate of the overall intensity of sexual selection across all lineages remains uncertain, not least because it focuses on a single sexual signalling modality, whereas many taxa engage in multimodal signalling [5]. Indeed, if the intensity of sexual selection targeted at one signalling modality (e.g. visual signals) trades off or is negatively correlated with the intensity of sexual selection targeted at another (e.g. acoustic signals) [3, 20], then such interactions could lead to a breakdown in the relationship between the underlying intensity of sexual selection across species and the visual traits used as proxies for sexual selection, therefore obscuring the true relationship between sexual selection and diversification [2, 8].

Progress in resolving this question has been slow because previous studies investigating the macroevolutionary consequences of sexual selection have generally focused exclusively on visual signalling traits [2, 21], leaving open the possibility that comparisons across different sexual signalling modalities may reveal contrasting patterns. Furthermore, most studies have failed to address the role of other important selection pressures potentially shaping the evolution of signal phenotypes, such as habitat differences, ecological divergence and interspecific interactions [22, 23], and have typically focused on geographically, taxonomically and/or ecologically restricted datasets, rather than sampling more broadly across major clades.

Here we address these issues by compiling data for a global sample of 259 avian species pairs from 33 passerine families to test the relationship between visible plumage dichromatism—used as a standard proxy for sexual selection in birds and other animals [7]—and macroevolutionary divergence in the other major avian signalling modality: song. We focus on birds because they offer an unequivocal example of multimodal sexual signalling in which both traits—avian plumage colouration (a visual ornament) and song (an acoustic ornament)—are known to function in inter- and intra-sexual selection in many avian taxa [5, 24-27]. In addition, the availability of complementary species-level data on avian morphological traits, ecology, biogeography and phylogeny allows us to assess the importance of plumage dichromatism in relation to a suite of key variables known to influence patterns of signal evolution.

Our analyses can be divided into three stages. First, we use published song recordings to estimate the extent of song divergence within species pairs. Second, we assess the relationship between sexual dichromatism and degree of song divergence across pairs. Third, we use multiple regression combined with model averaging techniques to assess the relative association between dichromatism and song divergence in relation to other factors. If negative relationships in the intensity of sexual selection between signalling modalities shape patterns of signal evolution in birds, we expect to find a negative correlation between plumage dichromatism and song divergence across species pairs. Alternatively, if traits from multiple signalling modalities have reinforcing or independent functions [28], we expect to find a positive relationship, or no relationship, between plumage dichromatism and song divergence across pairs, respectively.

Methods

SPECIES SAMPLING AND PHYLOGENETIC FRAMEWORK

We used published molecular phylogenies to select a sample of passerine species pairs for which high quality song recordings were available [see 2, 8]. Each pair consisted of sister species, i.e. pairs of lineages that represent each other's closest relative. We note that a few of our study pairs contain species that are not true sisters, both because of incomplete sampling in published phylogenies, and because we included some near-sisters in which one member of the pair belonged to a sister clade (or both species from a polytomy). This approach is based on the assumption that comparisons between near-sisters are informative about phenotypic divergence during recent evolutionary history [2]. Overall, our sample contained 518 species from 259 species pairs (including 243 sister species and 13 near-sisters) widely distributed across the passerine radiation (Fig. S1). For full details, see the supplementary online material. To provide a phylogenetic framework for our analyses, we sampled 1000 molecular-only trees from www.birdtree.org [29], which were pruned to include only the species included in our dataset. We then used TreeAnnotator [30] to generate a maximum clade credibility (MCC) tree, which was then pruned so that each pair was represented by a single tip.

SONG DIVERGENCE

To quantify the extent of song divergence within species pairs, we downloaded songs for all species from the Macaulay Library of Natural Sounds (www.macaulaylibrary.org) and the online database

Xeno Canto (www.xeno-canto.org). We digitized sound files in Raven Pro v1.4 using standard settings, then measured seven key temporal and spectral traits that together capture important interspecific differences in overall signal structure [for full details see 31, 32]: (i) maximum frequency (kHz), (ii) minimum frequency (kHz), (iii) peak frequency (kHz; frequency in the signal with the greatest amplitude), (iv) bandwidth (kHz; maximum frequency minus minimum frequency), (v) signal duration (s), (vi) number of notes and (vii) pace (number of notes s^{-1}). For each species, at least three high-quality recordings were measured (mean 4.8 recordings per species), providing a total sample of 2476 songs. To reduce the dimensionality of the dataset, we conducted a principal components (PC) analysis on the covariance matrix of individual (log-transformed) song measurements. The first three PCs from this analysis accounted for over 83% of the variance in the original acoustic dataset, with each PC capturing a distinct component of overall signal structure (Table 1). Specifically, PC1 (41%) loaded heavily with variables related to song pitch, PC2 (24%) loaded heavily with variables related to song length, and PC3 (18%) primarily loaded with song pace (Table 1). We therefore interpreted these PCs as axes of variation in song pitch (PC1), length (PC2) and pace (PC3), respectively, with variation in overall song structure captured by position in this three-dimensional acoustic space (Fig. S2). Using these PCs, we estimated within-pair song disparity for all pairs as the Euclidean distance between species' mean PC scores in terms of overall song structure (PC1-3), and in terms of song pitch, length and pace separately (Fig. 1).

SEXUAL DICHROMATISM

To quantify the degree of sexual dichromatism within pairs, we used the species-level scores of dichromatism generated by [8] based on human observations of published illustrations (see justification below). Briefly, we used standard methodology [17, 33] to score the difference in plumage coloration between the sexes over five body regions (head, nape-rump-back, throat-belly, tail, and wings) for each species in our sample. Each region was scored separately using three scores: 0, no difference between the sexes; 1, difference between the sexes only in shade or intensity of colour; 2, difference in colour or pattern between the sexes. The dichromatism scores for all five body regions were then summed to give species-specific scores of plumage dichromatism on a scale from 0 (monochromatic) to 10 (maximum dichromatism).

Although it would be preferable to estimate dichromatism using more objective measurements of bird plumage colouration, such as those generated by a spectrophotometer, these types of measurements are currently unavailable for most species. However, the typically high correlation between spectrophotometric and human estimates of dichromatism indicate that human vision can provide a reliable estimate of avian colouration in general [34] and sexual dichromatism in particular [35, 36]. To test this assumption in our dataset, we used spectrophotometric measurements of plumage colouration from museum specimens described in [8] to calculate objective estimates of dichromatism for species and pairs common to both datasets, following methods outlined previously

[2, 8]. This resulted in dichromatism estimates for 264 species (132 pairs), which we compared to our scores based on handbook illustrations using Spearman's rank correlation. We found that estimates of dichromatism based on spec measurements and human vision were highly positively correlated (Fig. S3) across both species (Spearman's $r = 0.76$, $P < 0.001$) and species pairs (Spearman's $r = 0.79$, $P < 0.001$). Thus, we conclude that scores of dichromatism based on published illustrations are valid for our dataset, in line with previous studies [e.g. 19].

ADDITIONAL PREDICTORS OF SONG DIVERGENCE

To explore the role of other factors known to influence estimates of phenotypic (particularly song) divergence in birds, we collected data for a suite of additional explanatory variables including divergence time [37], life history and allometric effects [38], migration status [39], habitat [40, 41], breeding latitude and insularity [42, 43], interspecific interactions [44], niche divergence [45-47] and song learning [48]. See Appendix S1 for full details of methods, data and data sources.

STATISTICAL ANALYSES

Estimating phylogenetic signal of song divergence

We used a generalised least squares (GLS) approach to test the phylogenetic signal of song divergence in our dataset. This approach, implemented in the R package *caper* [49], estimates a maximum likelihood (ML) value for phylogenetic signal (λ) [50], which typically varies between zero (trait variance is independent of phylogeny) and one (trait variance follows a Brownian motion model of evolution). In the context of analysing song divergence, a value of $\lambda = 0$ indicates that extent of song divergence within pairs is random with respect to phylogeny, whereas a value of $\lambda = 1$ implies that closely related pairs have more similar levels of song disparity than would be expected by chance. We found that ML values of λ were zero for all four measures of song divergence, with values of $\lambda = 1$ (i.e. a Brownian motion model of evolution) significantly rejected in all cases (Table S1). Results were qualitatively similar for an alternative dataset corrected for observed levels of intraspecific variation (see below; Table S1), indicating that variation in the extent of within-pair song divergence in our dataset is unrelated to phylogeny. This allowed us to use non-phylogenetic regression techniques with more flexible error structures than currently possible in a statistical phylogenetic comparative framework, which was necessary for our dataset (see below).

Testing the relationship between predictors and extent of song divergence

To model the observed variation in estimates of within-pair song divergence, accounting for the right-skewed distribution of disparity estimates (Fig. S4), we used generalised linear models (GLMs) with a gamma error distribution and log link. Using this approach, we (i) examined the relationship between song disparity and degree of plumage dichromatism, (ii) tested for an interaction effect between dichromatism and habitat type (forest dependency) and (iii) assessed the combined influence of all predictor variables on the extent of song disparity using single and multi-predictor regression and Akaike information criterion-based model averaging [51] corrected for small sample sizes (AIC_c).

To perform model averaging, following [52] and [53], we fitted models encompassing all possible additive combinations of our predictor variables (see above), including a null (intercept-only) model, calculating the AIC_c score of each model. We then calculated the relative importance (RI) for each predictor variable as the sum of relative Akaike weights for models in which they appear. RI values scale from 0 to 1, where a variable with a score of 0 is associated with very low Akaike weights (i.e. low importance) and 1 is consistently associated with high weights (i.e. high importance). We also calculated model-averaged estimates of regression parameters and standard error values, calculated as the sum of the parameter estimates for each model including that predictor, multiplied by the relative Akaike weight of each of those models. To give further insight into the relative importance of predictor variables, we also identified the variables included in the top-ranked (i.e. best fitting) model in each case. We used this procedure to assess the effect of predictors on our response variables, which were overall song disparity (PC1-3), as well as disparity in pitch (PC1), length (PC2) and pace (PC3) separately. For multipredictor models, we restricted the dataset to include only those species pairs for which complete data for all predictors were available (246 of 259 pairs) and pseudo-R^2 values for GLMs were estimated using the method of [54]. Pair age, generation length, body mass disparity and beak disparity were ln-transformed to improve the normality of residuals and models were inspected to ensure they complied with modelling assumptions. To improve the interpretability of regression coefficients, predictor variables were centered and standardised prior to model fitting [55]. All analyses were conducted in R version 3.3.1 and model averaging was performed using the R package *MuMIn* [56].

Collinearity among predictor variables is a potential problem in multiple regression modelling [57]. However, methods based on the type of information theoretic-based model averaging approach we employ here are generally robust to collinearity [57, 58]. Nonetheless, we checked for collinearity among predictors by calculating Pearson correlation coefficients (r^2) for all pairs of variables. Only one pair (latitude and migration) showed evidence of strong collinearity ($r^2 = 0.75$; Table S2). However, we found no evidence of significant effects for either of these variables (see *Results*), indicating that collinearity among predictors is unlikely to explain our results.

Sensitivity analysis: accounting for intraspecific variation

Intraspecific trait variation can generate overinflated estimates of interspecific disparity, especially when sample sizes are low relative to levels of intraspecific variation and/or measurement error [59]. To test whether our results were sensitive to these issues, we repeated our main analyses using an alternative version of our dataset in which within-species song disparity estimates were corrected for observed levels of intraspecific variation using an approach based on ANOVA [42, 60].

Results

RELATIONSHIP BETWEEN DICHROMATISM AND EXTENT OF SONG DIVERGENCE

Our models revealed that plumage dichromatism was significantly negatively correlated with overall song divergence between species (Table S3). Species pairs with a greater degree of plumage dichromatism tended to have less divergent songs than more monochromatic species pairs (Fig. 2). Analysing patterns of divergence in each song trait separately revealed that the overall effect of dichromatism was primarily driven by significant negative relationships with divergence in song pitch (PC1) and length (PC2), with more marginal effects on song pace (PC3) (Table S3). Furthermore, including an interaction effect with forest dependency in these models revealed no significant statistical support for the hypothesis that the relationship between dichromatism and song divergence is mediated by variation in habitat type across taxa (Table S3). Rerunning analyses correcting for observed levels of intraspecific variation produced highly similar results (Table S3).

ADDITIONAL PREDICTORS OF SONG DIVERGENCE

Single predictor regression models focused on our additional predictors of song divergence identified several variables that were individually correlated with variation in song disparity across pairs (Fig. S5-S8). In terms of total song disparity (PC1-3), the strongest individual predictor was pair age (Table S4). Furthermore, variation in overall song disparity was also significantly correlated with disparity in beak morphology, with more marginal effects detected for several other variables, including forest dependency and mass disparity (Table S4). We also detected additional significant correlations between individual predictors and estimates of disparity in specific components of song structure (Table S5).

We then assessed the relative importance of all predictors using AIC_c model averaging techniques (Fig. 3). The best-supported predictor of total song disparity (PC1-3) was pair age (RI = 0.99), which exhibited a strong and highly significantly positive relationship with disparity (Table S6). However, even after accounting for this relationship, the negative effect of dichromatism remained strong (RI = 0.98) (Fig. 3). The AIC_c best model for total song disparity accounted for 17% of the total variation, and retained these two variables plus mean pair body mass, forest dependency and within-pair beak disparity mass as positive effects, and confamilial sympatry as a negative effect (Fig. 3; Table S6).

Analysing relationships in each component song trait separately revealed that predictor variables had contrasting effects depending on the axis considered (Fig. S9; Table S7). On the one hand, divergence in song pitch (PC1) was best predicted by significant effects of dichromatism (RI = 0.97), pair age (RI = 0.96), body mass (RI = 0.93), confamilial sympatry (RI = 0.92) and forest dependency (RI = 0.89), whereas divergence in song length (PC2) was best explained by significant effects of only pair age (RI = 0.92) and dichromatism (RI = 0.91). In contrast, the only significant predictor of divergence in song pace (PC3) was a positive effect of within-pair beak disparity (RI = 0.82). Overall, AIC_c top models for these variables accounted for 22%, 12% and 6% of the total variation in disparity in song pitch, length and pace, respectively (Table S8). Rerunning models accounting for intraspecific variation produced qualitatively similar results (Table S10-S11).

Discussion

Our analyses reveal that the degree of sexual dichromatism is negatively related to the extent of divergence in song structure among closely related bird species, a pattern that remained strong after accounting for a suite of potentially correlated or confounding variables, as well as for intraspecific trait variation. This contrasts with the findings of previous studies reporting evidence for significant positive correlations between indictors of sexual selection and signal divergence in birds, supporting the view that sexual selection can drive parallel divergence across multiple signals [2, 21]. However, these studies assessed patterns of divergence in visual signalling traits (i.e. plumage colouration), using proxies for the intensity of sexual selection derived from the same signalling modality (e.g. dichromatism). In contrast, we have focused across major avian signalling modalities, finding the opposite relationship: that dichromatism (a visual signal) is negatively associated with divergence in song (an acoustic signal). Our results are therefore incompatible with a straightforward interpretation of the sexual selection-diversification hypothesis, whereby higher levels of dichromatism indicating strong sexual selection should be related to greater levels of mating signal divergence regardless of signalling modality. Instead, our findings are consistent with the alternative view that negative interactions between alternative signalling modalities play an important role in shaping macroevolutionary patterns of signal evolution in birds.

One intuitive explanation for the negative correlation between plumage dichromatism and song divergence is that it reflects an underlying link between sexual selection and acoustic signal divergence in species that do not rely on visual signals. This makes sense because single-species studies have demonstrated an important role for female choice and/or male-male competition in shaping many aspects of avian acoustic signal design [24], and many avian taxa with drab or monochromatic plumage are known to possess highly elaborate acoustic signals which often provide the best means of differentiating among lineages (e.g. Old World leaf warblers; Phylloscopidae) [61]. Thus, increased sexual selection on acoustic traits relative to visual traits in monochromatic taxa provides a plausible explanation for a negative relationship between plumage dichromatism and song disparity at broad macroevolutionary scales.

A key challenge facing this interpretation is to explain why, within species, selection would favour signals from one rather than multiple signalling modalities, thus generating negative relationships across modalities at a macroevolutionary scale. It is possible that the relative costs and benefits of signalling via a given sensory modality are shaped by the prevailing environmental conditions [41], such that ecological differences among species should play a role in determining the relative prominence of one signal type over another [27]. In line with this idea, bird species inhabiting dense habitats such as thickets and reedbeds tend to have much more elaborate songs than visual signals. However, our models including forest dependency (i.e. relative strength of association with forest habitat) as an interaction term provided no support for the idea that the relationship between plumage dichromatism and song divergence is mediated by broad-scale habitat differences among taxa.

An alternative explanation is that our findings reflect the signature of evolutionary trade-offs between alternate signalling modalities. Under a resource- or cost-based trade-off scenario—such as that envisaged by Darwin [3] and later termed the ‘transfer hypothesis’ [20]—constraints on sexual selection within species make it costly for males to signal in (or females to choose between) multiple signalling modalities [62, 63], generating the potential for interspecific trade-offs in ornament elaboration (and diversification) between alternate signalling modalities [64]. However, this explanation relies on the assumption that investment in one signalling modality constrains investment in another, which is perhaps unlikely for avian plumage and song traits given the differences in how such signals are produced and displayed [27]. An alternative explanation therefore is that our results reflect the outcome of ‘redundancy’ among alternate signal types. Under a redundancy-based model, the spread of an attractive signal in one modality leads simultaneously to increased selection for detecting the novel signal and a weakening of selection for elaborate signals in alternate modalities, which occurs not because of costs associated with producing or maintaining multiple sexual signals, but because sexual selection on the latter trait is weak or non-existent, due to redundancy [65]. Such redundancy-based trade-offs can theoretically occur in the absence of any habitat differences among taxa, or resource limitation underlying the production of signalling traits. Thus, whether selection favours one signal type (e.g. song) over another (e.g. plumage) largely depends on which signal type evolved first, which may largely be due to historical contingencies [66-68].

CONTRIBUTORY FACTORS

In addition to variation in the strength and targets of sexual selection, our results support a role for several other factors in shaping patterns of acoustic signal divergence in birds. We found strong evidence for a positive relationship between species pair age and degree of song disparity, in line with previous studies [e.g. 2, 42, 60, 69], as well as the general consensus that patterns of phenotypic divergence are primarily dictated by the time available for trait differences to evolve [37, 70]. Body mass also emerged as a significant predictor of song divergence, in line with previous studies indicating positive relationships between body mass and patterns of signal evolution in birds [2, 19]. Furthermore, we found support for links between song divergence and both habitat and the degree of confamilial sympatry. First, we found evidence for increased pitch disparity in species pairs with higher levels of forest dependency, consistent with the idea of stronger (divergent) selection on acoustic traits in taxa signalling in densely vegetated habitats [40, 41]. Second, we found that pairs which co-occurred with a greater proportion of confamilial species had lower levels of song divergence than those with lower levels of overlap. This accords with the view that interactions among related species can constrain phenotypic divergence [71], in part because acoustic communities appear to ‘partition’ finite aspects of acoustic signalling space [31, 72-74]. Finally, we found that divergence in song pace was significantly positively correlated with disparity in species’ beak morphology, presumably because biophysical constraints on song production generate correlated evolution between songs and beaks [46, 75, 76]. Even when we accounted for these significant effects in statistical models, the strong negative association between song divergence and dichromatism was retained.

IMPLICATIONS FOR COMPARATIVE STUDIES

Our finding that dichromatism is negatively related to song divergence across a broad sample of avian species pairs has important implications for studies testing macroevolutionary hypotheses related to sexual selection. Most importantly, it implies that plumage dichromatism provides a relatively ineffective proxy for the intensity of sexual selection in taxa primarily using non-visual signals. This potential limitation of dichromatism has previously been proposed [2, 8] with reference to bird species such as the common nightingale (*Luscinia megarhynchos*), common whitethroat (*Sylvia communis*) and sedge warbler (*Acrocephalus schoenobaenus*), passerine species with largely monomorphic plumage colouration, elaborate song traits and strong sexual selection [77-79]. Our results provide broad-scale empirical support for this view, and indicate that dichromatism will underestimate sexual selection in these taxa, potentially being negatively related to the intensity of sexual selection in samples dominated by non-visual signallers. Thus, the underlying effect of sexual selection may often be obscured in comparative studies based solely on dichromatism, perhaps helping to explain the weak or non-existent correlations between dichromatism and speciation rates in birds and other taxa engaged in multimodal signalling [7-11].

CONCLUSIONS

Taken together, our findings are consistent with the view that sexual selection plays a major role in shaping sexual signal evolution, in conjunction with ecological factors [23, 80, 81]. However, whereas most previous studies have focused on a single signalling modality, we found evidence of a negative relationship between visual and acoustic signalling in birds, supporting the more general view that negative interactions between signalling modalities can explain general patterns of signal evolution [3, 20]. Not only do these results suggest that such ‘trade-offs’ are important in shaping phenotypic diversity, they also indicate that phylogenetic tests based on phenotypic metrics for the intensity of sexual selection will underestimate the association between sexual selection and diversification. We conclude that the rigour and accuracy of any comparative analysis testing the effects of sexual selection will be improved by considering phenotypic proxies for sexual selection that span all relevant signalling modalities, be they visual, acoustic, or olfactory.

Acknowledgements

We thank Angela Chira, Emma Hughes, Joseph Llanos and Gavin Thomas for helpful discussion and three anonymous reviewers for constructive comments on the manuscript. We thank Mark Adams and Hein van Grouw at Natural History Museum for logistical assistance and access to specimens. Collection of trait data was supported by Natural Environment Research Grant (NE/I028068/1) to JAT and NS.

Data accessibility

The full dataset has been uploaded to Dryad Data Repository, and is available via the link: https://datadryad.org/review?doi=doi:10.5061/dryad.b4p43t7.

References

1. Lande R. 1981 Models of speciation by sexual selection on polygenic traits. *Proc Natl Acad Sci USA* **78**(6 I), 3721-3725.
2. Seddon N., Botero C.A., Tobias J.A., Dunn P.O., Macgregor H.E., Rubenstein D.R., Uy J.A., Weir J.T., Whittingham L.A., Safran R.J. 2013 Sexual selection accelerates signal evolution during speciation in birds. *Proc R Soc London Ser B* **280**(1766), 20131065. (doi:10.1098/rspb.2013.1065).
3. Darwin C.R. 1871 *The descent of man, and selection in relation to sex*. London, UK, John Murray.
4. West-Eberhard M.J. 1983 Sexual selection, social competition, and speciation. *Q Rev Biol* **58**(2), 155-183.
5. Andersson M. 1994 *Sexual selection*. Princeton, Princeton University Press.
6. Panhuis T.M., Butlin R., Zuk M., Tregenza T. 2001 Sexual selection and speciation. *Trends Ecol Evol* **16**(7), 364-371.
7. Kraaijeveld K., Kraaijeveld-Smit F.J.L., Maan M.E. 2011 Sexual selection and speciation: the comparative evidence revisited. *Biol Rev Camb Philos Soc* **86**(2), 367-377. (doi:10.1111/j.1469-185X.2010.00150.x).
8. Cooney C.R., Tobias J.A., Weir J.T., Botero C.A., Seddon N. 2017 Sexual selection, speciation and constraints on geographical range overlap in birds. *Ecol Lett* **20**(7), 863-871. (doi:10.1111/ele.12780).
9. Phillimore A.B., Freckleton R.P., Orme C.D.L., Owens I.P.F. 2006 Ecology predicts large-scale patterns of phylogenetic diversification in birds. *Am Nat* **168**(2), 220-229.
10. Huang H., Rabosky D.L. 2014 Sexual selection and diversification: reexamining the correlation between dichromatism and speciation rate in birds. *Am Nat* **184**(5), E101-E114. (doi:10.1086/678054).
11. Morrow E.H., Pitcher T.E., Arnqvist G. 2003 No evidence that sexual selection is an 'engine of speciation' in birds. *Ecol Lett* **6**(3), 228-234.

12. Gage M.J., Parker G.A., Nylin S., Wiklund C. 2002 Sexual selection and speciation in mammals, butterflies and spiders. *Proc R Soc London Ser B* **269**(1507), 2309-2316. (doi:10.1098/rspb.2002.2154).

13. Ritchie M.G., Hamill R.M., Graves J.A., Magurran A.E., Webb S.A., Macías Garcia C. 2007 Sex and differentiation: population genetic divergence and sexual dimorphism in Mexican goodeid fish. *J Evol Biol* **20**(5), 2048-2055. (doi:10.1111/j.1420-9101.2007.01357.x).

14. Barraclough T.G., Harvey P.H., Nee S. 1995 Sexual selection and taxonomic diversity in passerine birds. *Proc R Soc London Ser B* **259**(1355), 211-215.

15. Owens I.P.F., Bennett P.M., Harvey P.H. 1999 Species richness among birds: body size, life history, sexual selection or ecology? *Proc R Soc London Ser B* **266**(1422), 933-939.

16. Dunn P.O., Whittingham L.A., Pitcher T.E. 2001 Mating systems, sperm competition, and the evolution of sexual dimorphism in birds. *Evolution* **55**(1), 161-175.

17. Owens I.P.F., Hartley I.R. 1998 Sexual dimorphism in birds: why are there so many different forms of dimorphism? *Proc R Soc London Ser B* **265**, 397-407.

18. Dunn P.O., Armenta J.K., Whittingham L.A. 2015 Natural and sexual selection act on different axes of variation in avian plumage color. *Science Advances* **1**, e1400155.

19. Dale J., Dey C.J., Delhey K., Kempenaers B., Valcu M. 2015 The effects of life history and sexual selection on male and female plumage colouration. *Nature* **527**(7578), 367-370. (doi:10.1038/nature15509).

20. Gilliard E.T. 1956 Bower ornamentation versus plumage characters in bower-birds. *Auk* **73**(3), 450-451.

21. Gomes A.C.R., Sorenson M.D., Cardoso G.C. 2016 Speciation is associated with changing ornamentation rather than stronger sexual selection. *Evolution* **70**(12), 2823-2838. (doi:10.1111/evo.13088).

22. Cuthill I.C., Allen W.L., Arbuckle K., Caspers B., Chaplin G., Hauber M.E., Hill G.E., Jablonski N.G., Jiggins C.D., Kelber A., et al. 2017 The biology of color. *Science* **357**(6350), eaan0221. (doi:10.1126/science.aan0221).

23. Wilkins M.R., Seddon N., Safran R.J. 2013 Evolutionary divergence in acoustic signals: causes and consequences. *Trends Ecol Evol* **28**(3), 156-166. (doi:10.1016/j.tree.2012.10.002).

24. Catchpole C.K., Slater P.J.B. 2008 *Bird song: biological themes and variations, 2nd edition*. Cambridge, UK, Cambridge University Press.

25. Collins S.A. 2004 Vocal flirting and fighting: the functions of birdsong. In *Nature's music: the science of birdsong* (eds. Marler P., Slabbekoorn H.), pp. 39-79. San Diego, CA, Elsevier Academic Press.

26. Slabbekoorn H.W. 2004 Singing in the wild: the ecology of birdsong. In *Nature's music: the science of birdsong* (eds. Marler P., Slabbekoorn H.W.), pp. 178-205. San Diego, CA, Elsevier Academic Press.

27. Price T.D. 2008 *Speciation in birds*. Greenwood Village, CO, Roberts and Co.

28. Møller A.P., Pomiankowski A. 1993 Why have birds got multiple sexual ornaments? *Behav Ecol Sociobiol* **32**(3), 167-176.

29. Jetz W., Thomas G.H., Joy J.B., Hartmann K., Mooers A.O. 2012 The global diversity of birds in space and time. *Nature* **491**, 444-448. (doi:10.1038/nature11631).

30. Drummond A.J., Suchard M.A., Xie D., Rambaut A. 2012 Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol Biol Evol* **29**(8), 1969-1973. (doi:10.1093/molbev/mss075).

31. Tobias J.A., Planqué R., Cram D.L., Seddon N. 2014 Species interactions and the structure of complex communication networks. *Proc Natl Acad Sci USA* **111**(3), 1020-1025.

32. Tobias J.A., Aben J., Brumfield R.T., Derryberry E.P., Halfwerk W., Slabbekoorn H., Seddon N. 2010 Song divergence by sensory drive in Amazonian birds. *Evolution* **64**(10), 2820-2839. (doi:10.1111/j.1558-5646.2010.01067.x).

33. Owens I.P.F., Bennett P.M. 1994 Mortality costs of parental care and sexual dimorphism in birds. *Proc R Soc London Ser B* **257**(1348), 1-8. (doi:10.1098/rspb.1994.0086).

34. Bergeron Z.T., Fuller R.C. 2018 Using human vision to detect variation in avian coloration: how bad is it? *Am Nat* **191**(2), 269-276. (doi:10.1086/695282).

35. Armenta J.K., Dunn P.O., Whittingham L.A. 2008 Quantifying avian sexual dichromatism: a comparison of methods. *J Exp Biol* **211**(15), 2423-2430. (doi:10.1242/jeb.013094).

36. Seddon N., Tobias J.A., Eaton M., Ödeen A. 2010 Human vision can provide a valid proxy for avian perception of sexual dichromatism. *Auk* **127**(2), 283-292. (doi:10.1525/auk.2009.09070).

37. Tobias J.A., Cornwallis C.K., Derryberry E.P., Claramunt S., Brumfield R.T., Seddon N. 2014 Species coexistence and the dynamics of phenotypic evolution in adaptive radiation. *Nature* **506**(7488), 359-363. (doi:10.1038/nature12874).

38. Ryan M.J., Brenowitz E.A. 1985 The role of body size, phylogeny, and ambient noise in the evolution of bird song. *Am Nat* **126**(1), 87-100.

39. Collins S.A., de Kort S.R., Pérez-Tris J., Tellería J.L. 2009 Migration strategy and divergent sexual selection on bird song. *Proc R Soc London Ser B* **276**(1656), 585-590. (doi:10.1098/rspb.2008.1011).

40. Morton E.S. 1975 Ecological sources of selection on avian sounds. *Am Nat* **109**(965), 17-24.

41. Endler J.A. 1992 Signals, signal conditions, and the direction of evolution. *Am Nat* **139**(Suppl.), S125-S153.

42. Weir J.T., Wheatcroft D. 2011 A latitudinal gradient in rates of evolution of avian syllable diversity and song length. *Proc R Soc London Ser B* **278**(1712), 1713-1720. (doi:10.1098/rspb.2010.2037).

43. Morinay J., Cardoso G.C., Doutrelant C., Covas R. 2013 The evolution of birdsong on islands. *Ecology and Evolution* **3**(16), 5127-5140. (doi:10.1002/ece3.864).

44. Pfennig D.W., Pfennig K.S. 2010 Character displacement and the origins of diversity. *Am Nat* **176**, S26-S44. (doi:10.1086/657056).

45. Podos J., Hendry A.P. 2006 The biomechanics of ecological speciation. In *Ecology and biomechanics: a mechanical approach to the ecology of animals and plants* (eds. Herrel A., Speck T., Rowe N.P.). Boca Raton, FL, CRC Press.

46. Podos J. 2001 Correlated evolution of morphology and vocal signal structure in Darwin's finches. *Nature* **409**, 185-188.

47. Lawson A.M., Weir J.T. 2014 Latitudinal gradients in climatic-niche evolution accelerate trait evolution at high latitudes. *Ecol Lett* **17**(11), 1427-1436. (doi:10.1111/ele.12346).

48. Lachlan R.F., Servedio M.R. 2004 Song learning accelerates allopatric speciation. *Evolution* **58**(9), 2049-2063.

49. Orme C.D.L., Freckleton R.P., Thomas G.H., Petzoldt T., Fritz S.A., Isaac N., Pearse W.D. 2013 Caper: comparative analyses of phylogenetics and evolution in R. R package version 0.5.2. http://cran.r-project.org/package=caper.

50. Pagel M. 1999 Inferring the historical patterns of biological evolution. *Nature* **401**(6756), 877-884.

51. Burnham K.P., Anderson D.R. 2002 *Model selection and multimodel inference: a practical information-theoretic approach*. New York, Springer.

52. Grueber C.E., Nakagawa S., Laws R.J., Jamieson I.G. 2011 Multimodel inference in ecology and evolution: challenges and solutions. *J Evol Biol* **24**(4), 699-711. (doi:10.1111/j.1420-9101.2010.02210.x).

53. Wagner C.E., Harmon L.J., Seehausen O. 2012 Ecological opportunity and sexual selection together predict adaptive radiation. *Nature* **487**, 366-370. (doi:10.1038/nature11144).

54. Nagelkerke N.J.D. 1991 A note on the general definition of the coefficient of determination. *Biometrika* **78**(3), 691-692.

55. Schielzeth H. 2010 Simple means to improve the interpretability of regression coefficients. *Methods in Ecology and Evolution* **1**(2), 103-113. (doi:10.1111/j.2041-210X.2010.00012.x).

56. Bartoń K. 2017 MuMIn: multi-model inference. R package, version 1.15.6. Available at: https://cran.r-project.org/web/packages/MuMIn/index.html.

57. Graham M.H. 2003 Confronting multicollinearity in ecological multiple regression. *Ecology* **84**(11), 2809-2815.

58. Freckleton R.P. 2011 Dealing with collinearity in behavioural and ecological data: model averaging and the problems of measurement error. *Behav Ecol Sociobiol* **65**(1), 91-101. (doi:10.1007/s00265-010-1045-6).

59. Silvestro D., Kostikova A., Litsios G., Pearman P.B., Salamin N. 2015 Measurement errors should always be incorporated in phylogenetic comparative analysis. *Methods in Ecology and Evolution* **6**, 340-346. (doi:10.1111/2041-210x.12337).

60. Weir J.T., Wheatcroft D.J., Price T.D. 2012 The role of ecological constraint in driving the evolution of avian song frequency across a latitudinal gradient. *Evolution* **66**(9), 2773-2783.

61. Tietze D.T., Martens J., Fischer B.S., Sun Y.H., Klussmann-Kolb A., Päckert M. 2015 Evolution of leaf warbler songs (Aves: Phylloscopidae). *Ecology and Evolution* **5**(3), 781-798. (doi:10.1002/ece3.1400).

62. Iwasa Y., Pomiankowski A. 1994 The evolution of mate preferences for multiple sexual ornaments. *Evolution* **48**(3), 853-867.

63. Schluter D., Price T. 1993 Honesty, perception and population divergence in sexually selected traits. *Proc R Soc London Ser B* **253**, 117-122.

64. Shutler D. 2011 Sexual selection: when to expect trade-offs. *Biol Lett* **7**(1), 101-104. (doi:10.1098/rsbl.2010.0531).

65. Agrawal A.A., Conner J.K., Rasmann S. 2010 Tradeoffs and negative correlations in evolutionary ecology. In *Evolution since Darwin: the first 150 years* (eds. Bell M.A., Futuyma D.J., Eanes W.F., Levinton J.S.). Sunderland, MA, Sinauer Associates.

66. Wischmann S., Floreano D., Keller L. 2012 Historical contingency affects signaling strategies and competitive abilities in evolving populations of simulated robots. *Proc Natl Acad Sci USA* **109**(3), 864-868. (doi:10.1073/pnas.1104267109).

67. Ord T.J., Charles G.K., Hofer R.K. 2011 The evolution of alternative adaptive strategies for effective communication in noisy environments. *Am Nat* **177**(1), 54-64. (doi:10.1086/657439).

68. Mani G.S., Clarke B.C. 1990 Mutational order: a major stochasitc process in evolution. *Proc R Soc London Ser B* **240**, 29-37.

69. Price J.J., Lanyon S.M. 2002 Reconstructing the evolution of complex bird song in the oropendolas. *Evolution* **56**(7), 1514-1529.

70. Uyeda J.C., Hansen T.F., Arnold S.J., Pienaar J. 2011 The million-year wait for macroevolutionary bursts. *Proc Natl Acad Sci USA* **108**(38), 15908-15913.

71. Grether G.F., Peiman K.S., Tobias J.A., Robinson B.W. 2017 Causes and consequences of behavioral interference between species. *Trends Ecol Evol* **32**(10), 760-772. (doi:10.1016/j.tree.2017.07.004).

72. Chek A.A., Bogart J.P., Lougheed S.C. 2003 Mating signal partitioning in multi-species assemblages: a null model test using frogs. *Ecol Lett* **6**, 235-247.

73. Seddon N. 2005 Ecological adaptation and species recognition drives vocal evolution in Neotropical suboscine birds. *Evolution* **59**(1), 200-215.

74. Grant B.R., Grant P.R. 2010 Songs of Darwin's finches diverge when a new species enters the community. *Proc Natl Acad Sci USA* **107**(47), 20156-20163. (doi:10.1073/pnas.1015115107).

75. Derryberry E.P., Seddon N., Claramunt S., Tobias J.A., Baker A., Aleixo A., Brumfield R.T. 2012 Correlated evolution of beak morphology and song in the neotropical woodcreeper radiation. *Evolution* **66**(9), 2784-2797. (doi:10.1111/j.1558-5646.2012.01642.x).

76. Podos J. 2004 Vocal mechanics in Darwin's finches: correlation of beak gape and song frequency. *J Exp Biol* **207**(4), 607-619. (doi:10.1242/jeb.00770).

77. Buchanan K.L., Catchpole C.K. 2000 Song as an indicator of male parental effort in the sedge warbler. *Proc R Soc London Ser B* **267**(1441), 321-326.

78. Halupka K., Boroweic M. 2006 Male whitethroats, *Sylvia communis*, advertise their future contributions to parental care. *Behaviour* **143**(1), 1-14.

79. Bartsch C., Weiss M., Kipper S. 2015 Multiple song features are related to paternal effort in common nightingales. *BMC Evol Biol* **15**, 115. (doi:10.1186/s12862-015-0390-5).

80. Safran R.J., Scordato E.S., Symes L.B., Rodriguez R.L., Mendelson T.C. 2013 Contributions of natural and sexual selection to the evolution of premating reproductive isolation: a research agenda. *Trends Ecol Evol* **28**(11), 643-650. (doi:10.1016/j.tree.2013.08.004).

81. Mendelson T.C., Martin M.D., Flaxman S.M. 2014 Mutation-order divergence by sexual selection: diversification of sexual signals in similar environments as a first step in speciation. *Ecol Lett* **17**(9), 1053-1066. (doi:10.1111/ele.12313).

Figure and table captions

Figure 1. A phylogenetic tree of passerine species pairs (n = 259), showing within-pair disparity in overall song (PC1-3; innermost, black), pitch (PC2; red), length (PC3; blue) and pace (PC3; outermost, green). Size of points corresponds to relative within-pair song disparity. A version including species names is available in the supporting online material.

Figure 2. Scatterplot showing the relationship between total within-pair song disparity (PC1-3) and plumage dichromatism across 259 species pairs of passerine birds. Regression line (with prediction intervals, shaded) indicates the best-fitting relationship between the two variables.

Figure 3. Model averaged coefficient estimates from multipredictor GLMs predicting variation in within-pair song disparity (PC1-3) among passerine species pairs (n = 259). Points indicate the standardised effect sizes for each of the (scaled) predictor variables and lines indicate 95% confidence intervals (CI). Sizes of points represent the relative importance (RI) of each of the predictor variables, where a value of RI = 0 indicates low importance and a value of RI = 1 indicates high importance. Predictors included in the AIC_c top model are coloured (blue = negative effect; red = positive effect), with significant ($P < 0.05$) model-averaged coefficients shown in darker colours.

Table 1. Variable loadings and variance explained by the first three principal components (PCs) of an analysis of measurements of individual song recordings (n = 2476) for the species included in our analyses (n = 518). Standardised loadings of the main contributors to each component are highlighted in bold.