LETTER

The effects of life history and sexual selection on male and female plumage colouration

James Dale[1], Cody J. Dey[2], Kaspar Delhey[3,4], Bart Kempenaers[5] & Mihai Valcu[5]

Classical sexual selection theory[1–4] provides a well-supported conceptual framework for understanding the evolution and signalling function of male ornaments. It predicts that males obtain greater fitness benefits than females through multiple mating because sperm are cheaper to produce than eggs. Sexual selection should therefore lead to the evolution of male-biased secondary sexual characters. However, females of many species are also highly ornamented[5–7]. The view that this is due to a correlated genetic response to selection on males[1,8] was widely accepted as an explanation for female ornamentation for over 100 years[5] and current theoretical[9,10] and empirical[11–13] evidence suggests that genetic constraints can limit sex-specific trait evolution. Alternatively, female ornamentation can be the outcome of direct selection for signalling needs[7,14]. Since few studies have explored interspecific patterns of both male and female elaboration, our understanding of the evolution of animal ornamentation remains incomplete, especially over broad taxonomic scales. Here we use a new method to quantify plumage colour of all ~6,000 species of passerine birds to determine the main evolutionary drivers of ornamental colouration in both sexes. We found that conspecific male and female colour elaboration are strongly correlated, suggesting that evolutionary changes in one sex are constrained by changes in the other sex. Both sexes are more ornamented in larger species and in species living in tropical environments. Ornamentation in females (but not males) is increased in cooperative breeders—species in which female–female competition for reproductive opportunities and other resources related to breeding may be high[6]. Finally, strong sexual selection on males has antagonistic effects, causing an increase in male colouration but a considerably more pronounced reduction in female ornamentation. Our results indicate that although there may be genetic constraints to sexually independent colour evolution, both female and male ornamentation are strongly and often differentially related to morphological, social and life-history variables.

The extraordinary interspecific variation in bird colouration has provided model studies on animal ornamentation (Fig. 1). The many striking cases of sexual dichromatism in birds illustrate the power of sexual selection because such species often have highly polygynous mating systems[1,4]. However, other factors besides sexual selection can influence colour elaboration[15,16] and there is growing evidence that many female ornaments are adaptive and subject to direct selection[5,7]. Females often compete for ecological resources, and may use ornamental traits to mediate competitive interactions. Since male ornaments are also used during competition for non-sexual resources (for example, food, territories), ornament evolution in both sexes may be better understood through the concept of social selection[14], which encompasses both traditional sexual selection, and selection on non-sexual interactions[7,17]. Under this framework, male and female ornamentation should correlate with both sexual selection and life-history traits that

potentially influence the level of competition for resources. To date, the life-history traits associated with ornamentation in both males and females remain poorly understood.

A fundamental challenge in resolving these issues is the quantification of colour ornamentation in a way that allows meaningful interspecific comparisons. We developed a method to quantify colour elaboration by determining how 'male-like' a focal plumage is (Fig. 2). This approach, which can be used with any colour quantification technique (see Methods and Extended Data Fig. 1), has the important property of quantifying diverse colours using a single metric (that is, how male-like it is). Thus, birds with dramatically different appearances and of different sex can have similar scores (for example, Fig. 1a, d).
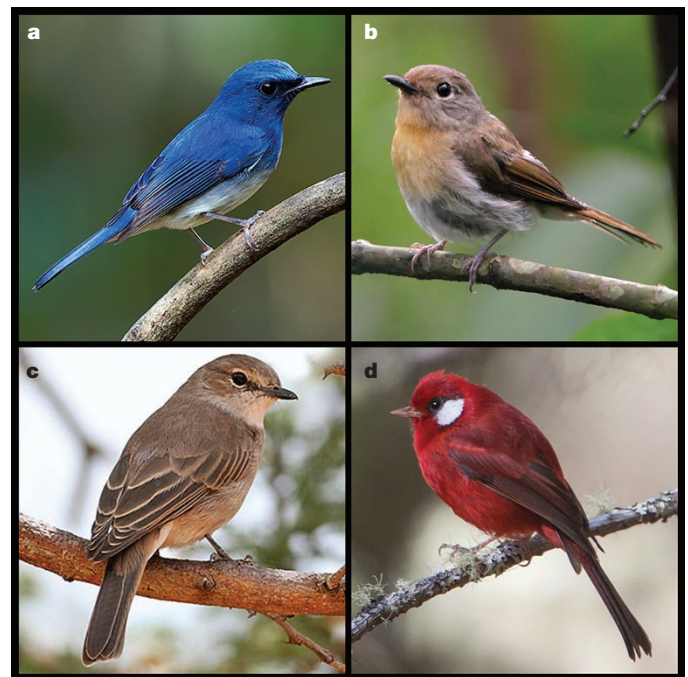

Figure 1 | Interspecific variation in avian plumage colouration.
a–d, Birds display an astonishing variety of colour patterns. The intense blue plumage of male Hainan blue flycatchers (*Cyornis hainanus*) (**a**) is dramatically more colourful than female plumage (**b**) in this strongly sexually dichromatic species. In contrast, pale flycatchers (*Bradornis pallidus*) are sexually monochromatic and both males (**c**) and females (not shown) are drab coloured. Other monochromatic species, however, can be highly ornamented, for example, female red warblers (*Cardellina ruber*) (**d**) have a degree of plumage colouration that rivals male Hainan blue flycatchers (plumage scores: 71.3 and 71.7, respectively). The evolutionary causes of this diversity are not well understood. Credits: **a**, S. Kongwithtaya; **b**, M. & P. Wong; **c**, M. Goodey; **d**, S. Colenutt.

[1]Institute of Natural & Mathematical Sciences, Massey University, Auckland 0745, New Zealand. [2]Department of Biology, McMaster University, 1280 Main St. West, Hamilton, Ontario L8S 4K1, Canada. [3]School of Biological Sciences, Monash University, Victoria 3800, Australia. [4]Max Planck Institute for Ornithology, Am Obstberg 1, 78315 Radolfzell, Germany. [5]Department of Behavioural Ecology and Evolutionary Genetics, Max Planck Institute for Ornithology, Eberhard Gwinner Str, 82319 Seewiesen, Germany.
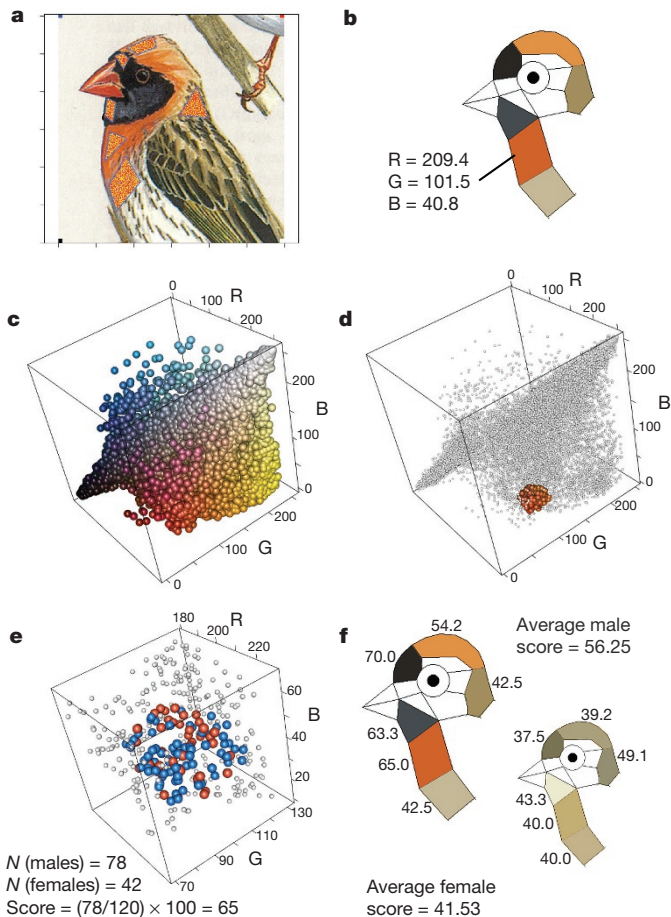
Figure 2 | The method used for quantifying plumage colouration.
a–**f**, For each sex of each species we identified similarly coloured plumage patches in all other species and calculated the proportion of those patches that are male. We illustrate this for the male red-billed quelea *Quelea quelea* (image from Lynx Edicions used with permission). **a**, Digitally scanned images of all passerine birds from the *Handbook of the Birds of the World*[18] ($N = 5,983$) were processed in the R package 'colorZapper' (see Methods). **b**, Red, green, blue (RGB) values in three dorsal (nape, crown, forehead) and three ventral (throat, upper breast, lower breast) patches were measured (RGB values for male upper breast shown). **c**, For each plumage patch, colour values for both sexes in all species were pooled. The panel depicts upper breast patch scores visualized in RGB colour space, $N = 11,966$ points (that is, 2 points for each species, point colour is determined from actual RGB values). **d**, **e**, For each sex in each species the nearest 1% ($N = 120$) of data points (in Euclidian space) were identified (coloured data points) (**d**) and the percentage of males (blue points) was determined and used as a 'patch score' (**e**). **f**, The average of the six patch scores was used as the final 'plumage score' for each sex (male and female red-billed quelea illustrated). Plumage scores thus reflect how 'male-like' or 'female-like' a plumage is, but are determined independently of the sex of each focal data point. There is a high correlation between scores determined as described here versus analogous scores determined with ultraviolet-to-visible reflectance spectra from museum specimens (Extended Data Fig. 1). The overall patterns reported in this paper are highly robust to different cut-offs used to calculate plumage scores (Extended Data Fig. 3).

We quantified male and female plumage colouration in all passerine birds (Order: Passeriformes) illustrated in the *Handbook of the Birds of the World*[18]. Passerines represent the most derived and largest avian radiation ($N = 5,983$ species, 61% of all birds). Figure 3a illustrates the basic patterns of colour variation in this group: in many species males are more colourful than females, but there are also many sexually monochromatic species with extensive variation in colour elaboration (Extended Data Fig. 2 shows the colours and associated scores for different patches).

Across the passerines, male and female plumage scores were highly correlated (Fig. 3a, phylogenetically controlled reduced major axis
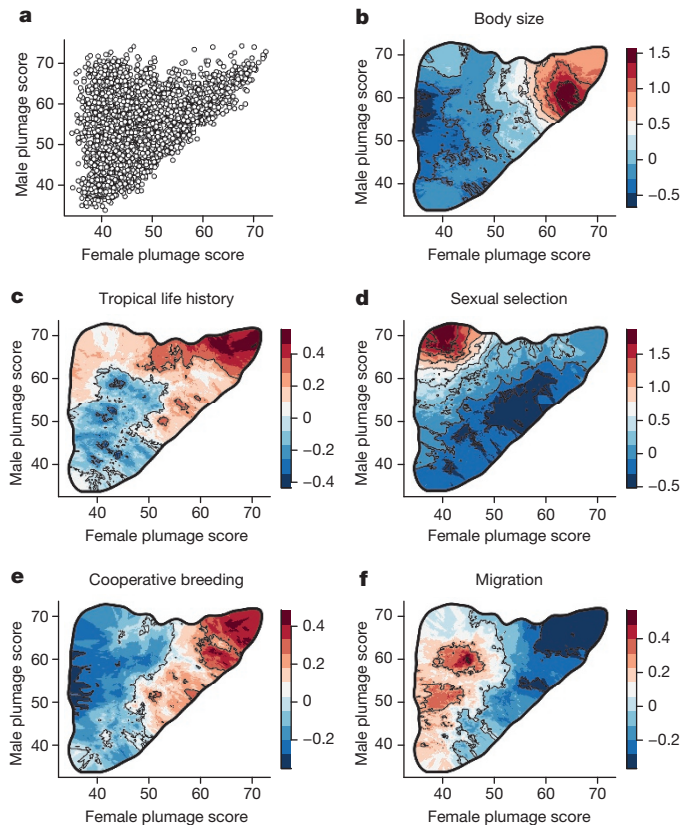

Figure 3 | Plumage scores and plumage dichromatism in relation to key predictors in passerine birds. a, Male plumage score versus female plumage score ($N = 5,983$ species). **b**–**f**, Contour maps depicting the average of the (scaled) predictor values overlaid within the 99.8% volume contour of the male versus female plumage score distribution. Note that the strength of the relationship between predictor values and plumage colouration varies between plots (as reflected by varying ranges in the contour legends). Values within the plots were calculated by superimposing a 300×300 grid over the scatter occurring between 30 and 80 and then calculating at each grid point the mean predictor value of the closest 3% of species.

regression, $R^2 = 0.299$, $T = 5.96$, $P < 0.0001$). Multivariate Ornstein–Uhlenbeck evolutionary models indicate that plumage colour evolution is subject to cross-sex constraints that partially restrict independent trait evolution (Extended Data Table 1). Cross-sex constraints, however, do not imply that ornamentation cannot also be directly selected for signalling needs[10].

To investigate colour diversity among species, we first tested for correlations between sexual dichromatism and ten predictor variables. Sexual dichromatism decreased with body mass and wing length, increased with latitude, more seasonal environments and clutch size, increased with social polygyny, sexual size dimorphism and female-only parental care, decreased with cooperative breeding and increased with migratory behaviour (Extended Data Table 2a).

Because many of the ten predictor variables are strongly intercorrelated, we next consolidated variation to five main predictors (Extended Data Table 3): (1) species with large 'body size' values were heavier and had longer wings; (2) species with high 'tropical life history' values were more likely to breed in the tropics, inhabit areas with year-long environmental stability and lay small clutches; (3) species with high 'sexual selection' scores tended to be socially polygynous, show male-biased sexual size dimorphism and lack paternal care; (4) 'cooperative breeding' was defined as present or absent; and (5) 'migration' as no, partial or complete migration between breeding and non-breeding ranges. Multiple predictor models (Extended Data Table 2b), which control for the confounding effects of the predictors, showed sexual selection to be the strongest predictor of dichromatism (followed by migration
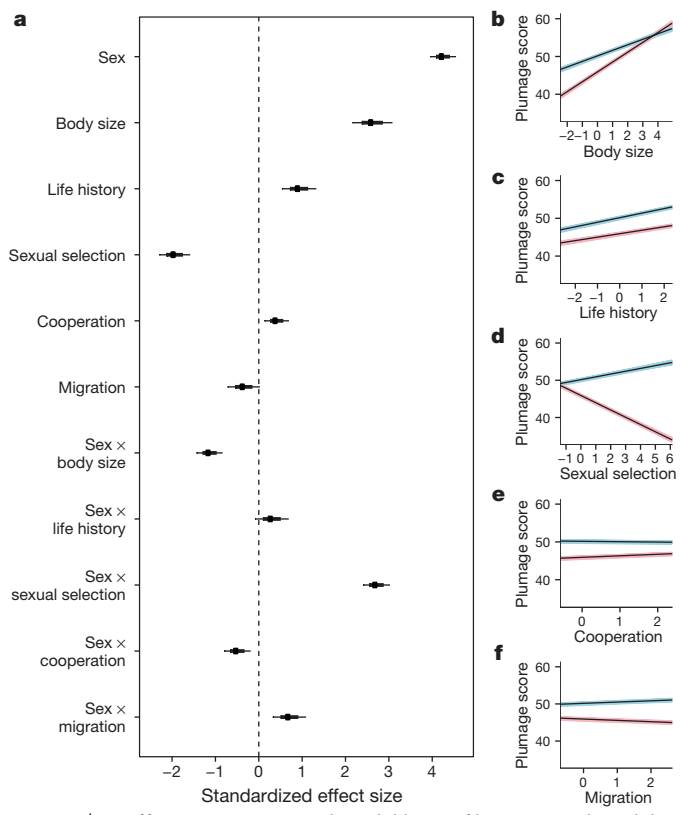
a

Figure 4 | Coefficient estimates and model lines of linear mixed models predicting plumage scores in 2,471 species of passerines. a, Coefficients plot of the effect sizes for each of the five key predictor variables (scaled) and their respective interactions with sex. The centre point denotes the mean, the thick bar denotes the posterior standard deviations (akin to the standard errors) and the thin bars denote the 95% lower and upper confidence limits as calculated by model-averaging 100 separate Markov chain Monte Carlo generalized linear mixed models using the package MCMCglmm[20] in R and 100 different phylogenetic trees from http://birdtree.org[30]. **b–f,** Model predictions of the effect sizes keeping the effects of other predictors constant (males, blue; females, pink). Solid black lines denote the mean effect size of the 100 separate MCMCglmm runs and coloured areas illustrate the full range of regression line estimates across all models.

and body size) while tropical life history did not have a significant effect. These results are difficult to interpret, however, because dichromatism can vary through evolutionary changes in males, females or both sexes[19]. Therefore, to resolve the evolutionary drivers associated with colour elaboration itself, we analysed male and female plumage phenotypes concurrently.

Phylogenetically informed linear mixed models[20] on complete data from 2,471 species (Fig. 4 and Extended Data Table 4) showed that the five predictor variables explained more interspecific variation in female than in male colouration. Female colouration was elaborated (more male-like) in larger species, in species with tropical life histories and in cooperative breeders. In males, larger species and species with tropical life histories also had more colour-elaborated plumage, but the effect of body size was smaller (Fig. 4b). As expected, in species with strong male-biased sexual selection, male colours were significantly more elaborated. However, sexual selection had a much larger and opposite effect on females: in species with strong male-biased sexual selection, females had greatly reduced colour elaboration (Fig. 4d). Plumage dichromatism therefore increased in species with strong sexual selection, as reported in numerous other studies[19], but our results show that the principal driver of this pattern is evolutionary change in female, not male, colour elaboration.

Because our predictor variables might interact in complex ways, we used phylogenetic path analysis[21] to disentangle cause and effect
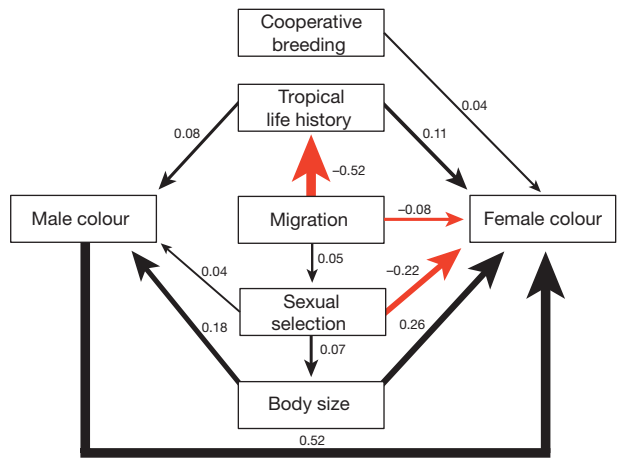

Figure 5 | Relationships among ecological variables and plumage colouration, as determined by phylogenetic controlled *d* separation path analysis[21]. Arrows indicate direct effects; the strength of the effect is indicated with numeric values and by line thickness. Arrow colour indicates the direction of the effect (black, positive; red, negative).

relationships. This supported the analysis described earlier, demonstrating that sexual selection has a strong direct negative effect on female colouration, and a weaker direct positive effect on male colouration (Fig. 5). Additionally, female colouration was directly influenced by all predictor variables. Models that included a correlated response to male colour were generally favoured, but models in which female colour was only directly influenced by male colour and not by other variables performed poorly (Extended Data Fig. 4 and Extended Data Table 5). Thus, our results suggest that female colouration is not merely a by-product of strong selection on males, but is an adaptive response to various social and life-history factors.

Three main conclusions follow from our study. First, although body size is rarely considered in interspecific studies of animal colouration (but see ref. 22), it strongly predicted colour elaboration in both sexes, and more strongly in females than in males. Larger species are both more colour elaborated and less dichromatic (Fig. 4b), consistent with the hypothesis that being larger reduces predation risk[23], thereby potentially weakening selection for crypsis. The result refutes the argument that large body size itself is an evolutionary constraint on colouration[22] (at least for passerines).

Second, both sexes of species with tropical life histories, typified by equatorial breeding ranges, low seasonality and small clutch sizes[24], were more elaborated than temperate breeding species (Fig. 4c), a robust verification of the hypothesis that tropical species are more colourful[25]. The strength of this effect rivals the strength of the effect of sexual selection on colouration (at least for males). These patterns are consistent with two non-mutually exclusive hypotheses. First, colourful plumage functions in mate choice and mutual sexual selection is stronger in tropical species; and second, colourful plumage functions in an aggressive context and resource competition is stronger in tropical species. Indeed, tropical species are thought to be under increased competition for breeding vacancies and resources, have longer-term pair bonds, more common year-round territoriality and increased convergence of male and female reproductive roles[24]. Moreover, there is an obvious convergence between visual and vocal signalling in tropical birds. Female song is more common in tropical species[24] and experimental evidence suggests that song in tropical species has dual functionality as both advertisement for mates as well as an 'armament' in competitive interactions[26].

Third, the intensity of sexual selection strongly predicted variation in plumage colouration (Fig. 4d): more elaborated males and increased sexual dichromatism are found in species with male-biased sexual selection. Increased sexual dichromatism, however, was driven mainly by a strong negative relationship between the intensity of sexual

selection and female colouration. A similar pattern was observed in New World blackbirds (Icteridae)[27], but our study demonstrates its generality within the passerines. This finding is consistent with two hypotheses. First, it can reflect increased sexual selection on females in monogamous species, supporting a game-theoretic model that predicts stable mutual mate choice in species with extensive parental investment in both sexes[28]. Second, selection for social signalling may be reduced in females of species with strong male-biased sexual selection[27]. Indeed, ecological factors that favour social polygyny[29] (for example, spatially clumped resources) may also be associated with reduced social competition between females, at least in the absence of paternal care. Additionally, high levels of male-biased sexual selection seem to break the correlation between male and female ornamentation (Fig. 4d). Here, the divergent parental roles of males and females may facilitate the evolution of sex-specific developmental modifiers (for example, hormones) that limit the expression of ornamental plumage in females[10].

Traditionally, studies have focused on male colour elaboration, particularly in species with extreme sexual selection. This has left much of the interspecific variation in colouration unexplained. Our results demonstrate clearly that the immense diversity in avian plumage colouration is the outcome of selection acting additively and often differentially between the sexes (Fig. 3). The patterns presented here for the passerines can be tested in other taxa and provide a rich arena for future hypothesis testing on the function of ornamentation in males and females.

1. Darwin, C. *The Descent of Man, and Selection in Relation to Sex* (John Murray, 1871).
2. Bateman, A. J. Intra-sexual selection in *Drosophila. Heredity* **2,** 349–368 (1948).
3. Trivers, R. in *Sexual Selection and the Descent of Man 1871–1971* (ed. Campbell, B.) 136–179 (Aldine, 1972).
4. Andersson, M. B. *Sexual Selection* (Princeton Univ. Press, 1994).
5. Amundsen, T. Why are female birds ornamented? *Trends Ecol. Evol.* **15,** 149–155 (2000).
6. Rubenstein, D. R. & Lovette, I. J. Reproductive skew and selection on female ornamentation in social species. *Nature* **462,** 786–789 (2009).
7. Tobias, J. A., Montgomerie, R. & Lyon, B. E. The evolution of female ornaments and weaponry: social selection, sexual selection and ecological competition. *Phil. Trans. R. Soc. B.* **367,** 2274–2293 (2012).
8. Lande, R. in *Sexual Selection: Testing the Alternatives* (eds Bradbury, J. W. & Andersson, M.) 83–94 (Wiley, 1987).
9. Bonduriansky, R. & Chenoweth, S. F. Intralocus sexual conflict. *Trends Ecol. Evol.* **24,** 280–288 (2009).
10. Kraaijeveld, K. Reversible trait loss: the genetic architecture of female ornaments. *Annu. Rev. Ecol. Evol. Syst.* **45,** 159–177 (2014).
11. Poissant, J., Wilson, A. J. & Coltman, D. W. Sex-specific genetic variance and the evolution of sexual dimorphism: a systematic review of cross-sex genetic correlations. *Evolution* **64,** 97–107 (2010).
12. Potti, J. & Canal, D. Heritability and genetic correlation between the sexes in a songbird sexual ornament. *Heredity* **106,** 945–954 (2011).
13. Cardoso, G. C. & Mota, P. G. Evolution of female carotenoid colouration by sexual constraint in Carduelis finches. *BMC Evol. Biol.* **10,** 82 (2010).
14. West-Eberhard, M. J. Sexual selection, social competition, and speciation. *Q. Rev. Biol.* **58,** 155–183 (1983).
15. Lyon, B. E., Eadie, J. M. & Hamilton, L. D. Parental choice selects for ornamental plumage in American coot chicks. *Nature* **371,** 240–243 (1994).
16. West-Eberhard, M. J. Darwin's forgotten idea: the social essence of sexual selection. *Neurosci. Biobehav. Rev.* **46,** 501–508 (2014).
17. Lyon, B. E. & Montgomerie, R. Sexual selection is a form of social selection. *Phil. Trans. R. Soc. B.* **367,** 2266–2273 (2012).
18. del Hoyo, J., Elliott, A. & Christie, D. A. *Handbook of the Birds of the World* Vols 8–16 (Lynx Edicions, 2003–2011).
19. Badyaev, A. V. & Hill, G. E. Avian sexual dichromatism in relation to phylogeny and ecology. *Annu. Rev. Ecol. Evol. Syst.* **34,** 27–49 (2003).
20. Hadfield, J. D. MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. *J. Stat. Softw.* **33,** 1–22 (2010).
21. von Hardenberg, A. & Gonzalez-Voyer, A. Disentangling evolutionary cause-effect relationships with phylogenetic confirmatory path analysis. *Evolution* **67,** 378–387 (2013).
22. Galván, I., Negro, J. J., Rodríguez, A. & Carrascal, L. M. On showy dwarfs and sober giants: body size as a constraint for the evolution of bird plumage colouration. *Acta Ornithol.* **48,** 65–80 (2013).
23. Ricklefs, R. E. Insights from comparative analyses of aging in birds and mammals. *Aging Cell* **9,** 273–284 (2010).
24. Stutchbury, B. J. & Morton, E. S. *Behavioral Ecology of Tropical Birds* (Academic, 2001).
25. Bailey, S. F. Latitudinal gradients in colors and patterns of passerine birds. *Condor* **80,** 372–381 (1978).
26. Tobias, J. A., Gamarra-Toledo, V., García-Olaechea, D., Pulgarín, P. C. & Seddon, N. Year-round resource defence and the evolution of male and female song in suboscine birds: social armaments are mutual ornaments. *J. Evol. Biol.* **24,** 2118–2138 (2011).
27. Irwin, R. E. The evolution of plumage dichromatism in the New World blackbirds: social selection on female brightness. *Am. Nat.* **144,** 890–907 (1994).
28. Kokko, H. & Johnstone, R. A. Why is mutual mate choice not the norm? Operational sex ratios, sex roles and the evolution of sexually dimorphic and monomorphic signalling. *Phil. Trans. R. Soc. B.* **357,** 319–330 (2002).
29. Emlen, S. T. & Oring, L. W. Ecology, sexual selection, and the evolution of mating systems. *Science* **197,** 215–223 (1977).
30. Jetz, W., Thomas, G. H., Joy, J. B., Hartmann, K. & Mooers, A. O. The global diversity of birds in space and time. *Nature* **491,** 444–448 (2012).

Author Contributions Conceived of the study: J.D., M.V. and B.K.; collected the data: J.D., M.V., K.D. and C.J.D.; developed the methods: J.D. and M.V.; analysed the data: J.D. and C.J.D. with help from M.V. and K.D.; wrote the paper: J.D. and C.J.D. with input from the other authors.

Author Information Data sets have been deposited in the Dryad Digital Repository (http://dx.doi.org/10.5061/dryad.1rp0s). Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.D. (j.dale@massey.ac.nz).

METHODS

Passerine classification and species list. Plumage scores were calculated on each of the 5,983 species of passerines (Order: Passeriformes) included in the *Handbook of the Birds of the World* (volumes 8–16)[18]. Phylogenetically controlled statistical analyses were restricted to the 5,831 of these species that are also included in the avian phylogenies at http://www.birdtree.org[30]. No statistical methods were used to predetermine sample size.

Plumage scores. We digitally scanned images of each passerine species from the plates in the *Handbook of the Birds of the World*[18] into 300 dpi JPEGs using a Fuji Xerox ApeosPort-IV C5575 set to default scan settings. Each species was cropped out of the scans and used to measure the RGB (red, green, blue) values for six patches (nape, crown, forehead, throat, upper breast, and lower breast) on each sex of each species using the R package 'colorZapper'[31]. We scored these regions because (1) they are consistently illustrated clearly for each species in the plates (rumps are often not shown for instance), and (2) the anterior body region is unarguably a very important signalling region for birds in general. For each plumage patch, a polygon was subjectively selected that encompassed the typical colouration evident in that area of the bird's plumage (see Fig. 2a for an example with red-billed quelea[32]). We excluded any obvious areas of glare added by the plate artists. colorZapper then calculated the mean values for R, G and B (on scales of 0 to 255) for 400 randomly chosen pixels within the selected polygon. In cases where multiple subspecies were illustrated, we scored colouration in the nominate subspecies.

Males and females are usually illustrated with the same image in the *Handbook of the Birds of the World*[18] when they are sexually monochromatic ($N = 3,822$ species). In these cases, the same image was measured independently twice, once to obtain male RGB values and once to obtain female RGB values. This was required to maintain similar measuring error between both dichromatic and monochromatic species. We confirmed that within-species measuring error was consistent between species scored with two different images versus species scored with the same image twice. In 63 of the 2,161 species where we used separate images for males and female colour measurements, females differed from males in ways other than the colour of the six plumage patches used in this study (for example, they had shorter tails or a different coloured iris). The correlation between male and female plumage scores from this sample ($R^2 = 0.905$, $N = 63$) was similar to the correlation between male and female plumage scores measured separately from the same image ($R^2 = 0.913$, $N = 3,822$).

Sometimes, when there is only a small difference between male and female plumages, the same image is used to illustrate both sexes of a species and the difference between the sexes is described in the text. In 372 species (6.22% of total) we noted a described difference in the text that was not illustrated in the plates. Many of these descriptions noted that females were "similar to males" but "duller" ($N = 65$), "slightly duller" ($N = 44$), "paler" ($N = 46$), or "slightly paler" ($N = 29$). Most of the rest of the described differences related to specific colour differences (for example, females described as "browner", "more rufous", "less blue", and so on) or to specific colour patches (for example, "coronal patch absent or small"). To estimate the magnitude of the measurement error associated with this issue, we identified 100 species for which the described difference between the sexes was similar to those described earlier (that is, "duller", "slightly duller", "paler", or "slightly paler") but for which we scored male and female colouration from separate images. In these species, the average difference between male and female plumage scores was 3.82 ($N = 100$ species, $P < 0.001$).

Because these described differences occur in a small percentage of species and reflect relatively small differences between the sexes, we expected them only negligibly to affect our general results. To confirm that our main results are robust to the error generated by these species, we repeated the main MCMCglmm analysis in two different ways: (1) with these species removed entirely from the data set (note that these species affected 168 of the 2,471 species (6.8%) in the MCMCglmm); and (2) with these species' plumage scores adjusted by subtracting the mean male–female difference calculated as described earlier (that is, 3.82) from the female scores. Both of these analyses yielded essentially identical results to the analysis reported in the paper (the R^2 between the effects of these analyses and the effects reported in Extended Data Table 4 was 0.998 and 0.999, respectively).

Scoring colouration with handbook plates represents a valid alternative to measuring colour on all the world's passerines using museum specimens and/or live individuals. Indeed, the objective of handbook plates is to reflect as accurately as possible the typical colouration and patterning of a species such that the image is a suitable reference for field identification. As such, special care is taken to reproduce colouration accurately and the plates in the *Handbooks of the Birds of the World* are highly regarded as superb in quality and consistency (for example, see ref. 33). Moreover, previous studies have found that colouration in plates is highly correlated with colouration in museum specimens (as measured by spectrometry[34]), and several previous comparative studies have used colouration in plates to test hypotheses about avian colour evolution (for example, see refs 35–38). Finally,

we validated that plumage scores generated with handbook plates were highly correlated with plumage scores generated with ultraviolet-to-visible (UV–Vis) spectrometry (see later and Extended Data Fig. 1).

Plumage scores validation analysis. We validated that plumage scores are consistent between different colour-measuring methodologies. We used UV–Vis spectrometry to measure the reflectance of 8 plumage patches (upper back, dorsal neck, crown, forehead, throat, upper breast and lower breast) of up to 3 male and 3 female museum specimens for 534 species of Australian terrestrial birds (229 non-passerines and 305 passerines). Museum specimens were obtained from Melbourne Museum (Museum Victoria) and the Australian National Wildlife Collection (CSIRO). Reflectance spectra were collected using a spectrometer (Avaspec 2048, Avantes) connected to a xenon pulsed light source (Avalight-XE) through a fibre optics cable fitted at the end with a plastic cylinder to exclude ambient light and standardize measuring distance. Reflectance spectra were expressed relative to a WS-2 white standard using Avasoft software. For each species we calculated mean reflectance spectra per patch separately for males and females. To compute plumage scores we pooled together all male and female mean spectra for each patch ($N = 1,068$ spectra for each patch). We matched each spectrum in the pool to the 120 most similar spectra as determined by minimizing the squared-differences between spectra at 5 nm wavelength intervals, summed across 300–700 nm. We then quantified the percentage of those 120 'closest matches' that corresponded to male spectra. For each species, male and female UV–Vis plumage scores were calculated as the mean 'percentage male' values across all of the eight patches measured in each sex of each species. For these same 534 species we also quantified patch colour using handbook plates digitally scanned from those published previously[18] and then scoring them in RGB colour space using the R-package colorZapper[31] (see earlier and Fig. 2). The 120 closest matching colour patches to any focal patch were determined by minimizing the Euclidean distance in RGB space, and then the proportion males in those closest matches was quantified. Note that this method of quantifying colouration is sample-size dependent: larger samples of species will provide more accurate measures of how 'male-like' any kind of colour is. Nevertheless, despite having only 534 species in this sample (contrasting with 5,983 species in the main analysis), both 'male-like' indices were strongly positively correlated ($R^2 = 0.67$, $P < 0.0001$; Extended Data Fig. 1a), indicating that using handbook plates to estimate bird colours is a suitable alternative to the use of reflectance spectrometry.

To demonstrate further that plumage scores calculated with handbook plate measurements are suitably interchangeable with plumage scores calculated with UV–Vis spectra, we repeated statistical analyses presented in the paper with the subset of passerine species for which spectral data were collected. If handbook plates provide a suitable surrogate measure then conclusions drawn from spectral data on the 305 species should be similar to conclusions drawn from the same 305 species based on plate data. First, we compared effect size estimates for single predictor PGLS models run on each of the ten predictor traits in Extended Data Table 2a. We ran models predicting female scores, male scores and dichromatism scores ($N = 30$ effects in total). Our analysis showed a very strong correlation between RGB effect sizes versus UV–Vis effect sizes (Extended Data Fig. 1b). Second, we compared effect size estimates (including sex–predictor interactions) for MCMCglmm models run with our five main predictors in Fig. 4. For each plumage score type we ran five models using a different phylogenetic tree[30] in each model. This analysis showed a very strong correlation between RGB effect sizes versus UV–Vis effect sizes (Extended Data Fig. 1c). In combination, these analyses strongly indicate that the biological patterns reported in the main analysis would not have been different if we had used plumage scores based on UV–Vis spectra. These results thus provide critical validation of our method because although human and avian vision have considerable overlap[19,39], birds can also see UV light not visible to humans[40].

Predictor variables. Body size (mass and wing length) was tested because it is a known confound of sexual size dimorphism[41] and life-history traits[42]. In addition, based on a previous study[22] we predicted a negative relationship between body size and colour elaboration. Tropical life history (latitude, seasonality and clutch size) was tested to evaluate the prediction that characteristics associated with tropical breeding—in particular, density-dependent or 'K-selected'[43] factors such as increased competition for limited breeding vacancies[24]—were associated with increased plumage colour elaboration. Sexual selection (social mating system, sexual size dimorphism, and paternal care) was tested because the prevailing view in the literature is that sexual selection on males is the fundamental evolutionary driver of male colour elaboration and sexual dimorphism. However, recent studies on New World blackbirds[27,44,45] (Icteridae), tanagers[46] (Thraupidae) and fairy-wrens[47] (Maluridae) have demonstrated rapid evolutionary transitions in female colouration (also see ref. 48). Therefore we also expected sexual selection on males to potentially have a strong effect on female colouration. Cooperative breeding species were predicted to be more colour elaborated because of increased social

competition among group members and this effect was predicted to be stronger in females than in males[6,49]. Migration was tested to evaluate two opposing predictions. On the one hand, it has been argued[50] that species with long-distance migration should be less colourful as a result of increased predation associated with long-distance movements (also see ref. 51). In contrast, other researchers[19,52] have argued that migratory species should be more colourful because migration imposes a short mate sampling period and enhances selection on signals related to migration ability.

Body mass data was taken from that described previously[53]. When more than one body mass entry was available for a species, we computed the mean weighted by sample size. Log-transformed values were used in the statistical analysis.

Wing length data was taken from that described previously[41]. Up to 7 different sets (average = 2.5) of wing measurements were recorded per species. Each set comprised the means (or mid-ranges when only a range was provided) for both males and females measured in a single population, and we took the means of these for final species values. Log-transformed values were used in the statistical analysis.

For latitude data, species' geographical location was computed based on the breeding range maps of all species[54]. First, breeding range polygons were transformed from an unprojected coordinate system (latitude, longitude) to an equal-area projected coordinate system (cylindrical equal-area, latitude of the origin: 0°). Species' geographical location was then computed as the latitude (degrees from equator) of the breeding range centroid.

For seasonality, Moderate Resolution Imaging Spectroradiometer (MODIS) land surface temperature rasters (code name MOD11C2) were obtained through the http://reverb.echo.nasa.gov gateway at a 0.05° spatial resolution and an 8-day temporal resolution (time span: 2000–2012)[55]. All the raster files were then superposed and a temperature time series was obtained for each pixel[54]. A coefficient of variation (CV%) of temperature was computed for each pixel[54]. For each species the CV% for all pixels within its breeding range were extracted. A species 'seasonality' score was defined as the median CV% score within its breeding range (log transformed).

Clutch size was compiled from standard references, in particular[18,56–60]. Up to five different reports of clutch sizes were recorded for each species, with the mean of these taken as the final clutch size of a species. Clutch size was recorded as the mean clutch size or, when only range data was provided, as the mid-range value. Log-transformed values were used in the statistical analysis.

Sexual size dimorphism was calculated as log(male wing length) − log(female wing length)[41], providing a proportional index of relative sizes of the sexes. Positive values reflect species where males are larger than females.

Social polygyny was scored on a four-point scale[36], with 0 = strict social monogamy (for example, zebra finch *Taeniopygia guttata*), 1 = monogamy with infrequent instances of polygyny observed (<5% of males, for example, lazuli bunting *Passerina amoena*), 2 = mostly social monogamy with regular occurrences of facultative social polygyny (5 to 20% of males, for example, American redstart *Setophaga ruticilla*), and 3 = obligate resource defence polygyny (>20% of males, for example, red-winged blackbird *Agelaius phoeniceus*) or lek polygyny (for example, lance-tailed manakin *Chiroxiphia lanceolata*). Assignments were made based on standard references, in particular[18,41,56–62]. There are a small number of passerine species with polygynandrous mating systems (for example, the dunnock *Prunella modularis*, Smith's longspur *Calcarius pictus* and sickle-billed vanga *Falculea palliata*) and these species were pooled with the monogamous species. We reasoned that sexual selection would be more similar in each sex in polygynandrous species compared with polygynous species, and our social polygyny scores were intended to specifically quantify male-biased sexual selection.

Paternal care was scored as absent (0) or present (1) primarily based on the data provided in ref. 63, including both the known and inferred data categories. For species in our data set not present in ref. 63, we used standard references, in particular ref. 18, to obtain the additional parental care scores.

Cooperative breeding was scored as absent (0), suspected (0.5), or present (1) primarily based on the data provided in ref. 63, including both the known and inferred data categories. For species in our data set not present in ref. 63, we used standard references, in particular ref. 18, to obtain the additional cooperative breeding scores.

Migration was scored on a scale from 0 to 2, with 0 = resident (that is, breeding and non-breeding ranges identical), 1 = partial migration (that is, some overlap between breeding and non-breeding ranges), 2 = complete migration (that is, no overlap between breeding and non-breeding ranges). Assignments were made based on the range maps published previously[18]. The migratory behaviour of one species, the Red Sea Swallow *Hirundo perdita*, is unknown.

Statistical analysis. Because closely related species tend to be more similar to each other than distantly related species, we used phylogenetically informed comparative analyses to account for potential non-independence among species owing to common ancestry. More specifically, the error structure of the statistical model

incorporates the degree of non-independence between species as estimated from the phylogeny. Phylogenetically informed methods are unlike ordinary statistical models (where data points are assumed to be independent) because they explicitly model how the covariance between species declines as they become more distantly related[64–66]. We used the Hackett[67] backbone phylogenetic trees available at http://birdtree.org[30] to estimate phylogenetic separation in our statistical models.

The relationships between the ten predictor variables listed earlier and sexual dichromatism (Extended Data Table 2) were modelled with phylogenetic generalized least-squares[64–66,68] (PGLS) using the R-packages 'ape'[69] and 'nlme'[70].

To calculate phylogenetic reduced major axis regression[71] and the phylogenetic principal components[72] used in the multiple predictor PGLS, MCMCglmm and paths analysis we used the R-package 'phytools'[73].

All data were analysed in R[74] version 3.1.0. To improve the interpretability of regression coefficients[75], predictor variables were centred and standardized to a mean = 0 and standard deviation = 1.

Phylogenetically informed generalized linear mixed models. Monte Carlo Markov chain generalized linear mixed models (Fig. 4 and Extended Data Table 4) were generated with the R-package 'MCMCglmm'[20]. MCMCglmm is a Markov chain Monte Carlo sampler for multivariate mixed models that enables the inclusion of a phylogeny as a design matrix in a Bayesian generalized linear modelling framework. The design matrix for phylogenetic effects was based on 100 trees from http://birdtree.org[30]. We then used MCMCglmm to fit male and female plumage scores as the response and the five predictor variables as continuous predictors. Phylogenetic effects were considered a random effect, sex was fit as a dummy variable and species was fit as an observation level random effect. We used the prior: [list(R = list(V = 1, nu = 0.002), G = list(G1 = list(V = 1, nu = 0.002)))] and model outcomes were insensitive to prior parameterization. We let the MCMC algorithm run for 10,000,000 iterations, with a burn in period of 3,000 and a sampling interval of 10,000. Each model generated ~1,000 independent samples of model parameters (Extended Data Table 4). Independency of samples in the Markov chain was assessed by graphic diagnostics and testing for autocorrelation between samples.

Multivariate Ornstein–Uhlenbeck evolutionary models. Ornstein–Uhlenbeck (OU) models are powerful tools for analysing the evolution of traits, and are well suited to studies of correlated evolution[76]. These models are a generalization of Brownian motion (BM) models[77], in that they incorporate stochastic trait evolution (defined by a drift parameter σ), but also allow trait values to be attracted towards optima, at a rate that is dependent on the value of α (the strength of selection towards the optima). When $\alpha = 0$, the process reduces to Brownian motion; however, many patterns of trait evolution are better fit by OU models, because trait evolution is often subject to stabilizing selection, at least over evolutionary time scales[78]. Additionally, the development of multivariate OU (mvOU) models has allowed trait evolution to be influenced by interactions with other trait values. Indeed, trait evolution is often thought of as a multivariate process, and mvOU models allow researchers to explore the influence of co-adaptive or limiting forces on trait evolution.

We used mvOU models to test whether the evolution of plumage ornamentation was subject to cross-sex constraints. Using the mvMORPH[79] package in R, we constructed five evolutionary models of the potential relationship between male and female ornamentation (using the plumage scores described in the main text ($N = 5,831$ species)) and compared their fit with AIC[76]. In model 1, male and female ornamentation evolved completely independently (that is, both the α and σ matrix were restricted to be diagonal) under OU processes, simulating no evolutionary relationship between male and female plumage colour. In model 2, male and female plumage scores evolved independently towards their optima (a diagonal α matrix); however, we allowed covariation in the stochastic element (that is, drift) of the OU process. This model might reflect a situation where male and female ornamentation respond similarly to some environmental parameter, but are not directly constrained by one another. In model 3, the evolution of male and female plumage ornamentation was influenced by an interaction with the other sex's trait value (a symmetric positive α matrix), and we allowed covariance in the stochastic element of trait evolution. Such a model represents cross-sex constraints in plumage colour evolution. Models 4 and 5 were BM models including independent and covarying drift matrices, respectively.

Model 3 was the clear best model (Extended Data Table 1a), demonstrating that male and female plumage colour do not evolve independently. Positive off-diagonal elements in the α matrix of this model (Extended Data Table 1b) suggest that male and female plumage values are pulled towards one another, consistent with cross-sex constraints against independent trait evolution, and the correlated response hypothesis. Models of fully independent trait evolution performed poorly (model 1 and 4), and models including covariation in drift performed better than those with independent drift. This suggests that in addition to an interaction in selection on male and female plumage colour, the sexes generally have a correlated response to ecological conditions.

Phylogenetic path analysis. Confirmatory path analysis is a special case of structural equation modelling used to build models of causal relationships among a suite of variables and test how the data conform with those causal models[80,81]. When conducted in a phylogenetic framework[21,82], this approach can infer the most likely evolutionary pathways by simultaneously considering both indirect and direct effects among variables, as well as the magnitude of these effects. This technique is advantageous in that it better reflects the true nature of evolutionary processes because factors do interact in complex ways (for example, by acting as both causal parents and causal children) to create evolutionary outcomes (that is, phenotypes). Here, we used the *d*-separation[21,80,82] method to test for the most likely relationships among seven variables related to plumage colouration. The variables used were the five principal components derived from the phylogenetic principle components analysis (Extended Data Table 3), as well as male plumage score and female plumage score. First, we built 14 biologically relevant models of the relationships among these variables (Extended Data Fig. 4). These models are constructed as directed acyclic graphs, which are required for this type of analysis. Model construction was based on the hypothesized relationships among variables that had been suggested in previous studies (for example, see refs 19, 27, 28) and from exploratory analysis of correlations among variables. Each model is then converted to a set of conditional independencies of the form (X1, X2) {X3}, where variables X1 and X2 are independent conditional on variable X3 (see elsewhere[80,82] for details). We then tested the set of conditional independencies in each model using phylogenetic generalized least squares (GLS) models (using the nlme[70] package in R[74]) and a single tree that was randomly selected from the tree pool at http://birdtree.org[30]. Using these GLS models, we calculated Fisher's *C*-statistic (which is equivalent to a maximum-likelihood estimate[83]) and the *C*-statistic information criterion (CICc) for each conceptual model[82,83].

Comparison of models based on CICc values indicated that model K was the best model (Extended Data Table 5). Model L appears to also be a competitive model, however, visual comparison of models K and L show that they suggest the same causal relationships among variables, except model L contains one additional causal link (between life history and sexual selection). The addition of a single, uninformative variable often causes marginal changes in CIC and can therefore appear to create competitive models, although in such a case the simpler model should be preferred[84,85]. Models including a direct effect of male colour on female colour performed better than those that did not include this effect (G, I) and those in which there was a direct effect of female colour on male colour (M, N). However, a model in which female colour was only directly influenced by male colour and not by other variables (C) also performed poorly. This suggests that female colouration is not simply a genetically correlated response to selection on male colouration, and is instead affected by other variables. The best model (K) generally supports the results of the MCMCglmm analysis (Fig. 4 and Extended Data Table 4). This model (Fig. 5) suggests that body size, cooperative breeding, migration, tropical life history, sexual selection and male colour all directly influence female colouration, while only sexual selection, tropical life history and body size directly influence male colouration. The magnitude and direction of the standardized regression coefficients (Fig. 5) suggest that there is a strong role of social selection on female colouration. Females become more colourful with increasing body size, more tropical life histories and cooperative breeding, factors that are predicted to increase the intensity of competition (see main text). Also consistent with our MCMCglmm analysis, the path analysis shows that sexual selection has direct, but antagonistic effects on both male and female colouration. Interestingly, the negative effect of sexual selection on female colouration is larger than the positive effect on male colouration. This pattern supports the finding that sexual selection has an overall negative effect on passerine colouration.

Code availability. Scripts of analyses and code used for figure production are available upon request from J.D.

31. Valcu, M. & Dale, J. colorZapper: color extraction utilities. R package version 1.0. https://github.com/valcu/colorZapper (2014).
32. Craig, A. in *Handbook of the Birds of the World* Vol. 15 (eds Del Hoyo, J., Elliot, A. & Christie, D. A.) (Lynx Edicions, 2010).
33. Starck, J. M. Review of *Handbook of the Birds of the World*. *Ethology* **102**, 436–440 (1996).
34. Badyaev, A. V. & Hill, G. E. Evolution of sexual dichromatism: contribution of carotenoid-versus melanin-based colouration. *Biol. J. Linn. Soc.* **69**, 153–172 (2000).
35. Gray, D. A. Carotenoids and sexual dichromatism in North American passerine birds. *Am. Nat.* **148**, 453–480 (1996).
36. Owens, I. P. F. & Hartley, I. R. Sexual dimorphism in birds: why are there so many different forms of dimorphism? *Proc. R. Soc. Lond. B* **265**, 397–407 (1998).
37. Olson, V. A. & Owens, I. P. F. Interspecific variation in the use of carotenoid-based colouration in birds: diet, life history and phylogeny. *J. Evol. Biol.* **18**, 1534–1546 (2005).
38. Dey, C. J., Valcu, M., Kempenaers, B. & Dale, J. Carotenoid-based bill coloration functions as a social, not sexual, signal in songbirds (Aves: Passeriformes). *J. Evol. Biol.* **28**, 250–258 (2015).
39. Seddon, N., Tobias, J. A., Eaton, M. & Odeen, A. Human vision can provide a valid proxy for avian perception of sexual dichromatism. *Auk* **127**, 283–292 (2010).
40. Cuthill, I. C. in *Bird Colouration, Volume 1: Mechanisms and Measurements* (eds Hill, G. E. & McGraw, K. J.) 3–40 (Harvard Univ. Press, 2006).
41. Dale, J. *et al.* Sexual selection explains Rensch's rule of allometry for sexual size dimorphism. *Proc. R. Soc. B* **274**, 2971–2979 (2007).
42. Calder, W. A. *Size, Function, and Life History* (Courier Corporation, 1996).
43. Pianka, E. R. On *r*- and *K*-selection. *Am. Nat.* **104**, 592–597 (1970).
44. Hofmann, C. M., Cronin, T. W. & Omland, K. E. Evolution of sexual dichromatism. 1. Convergent losses of elaborate female colouration in New World orioles (*Icterus* spp.). *Auk* **125**, 778–789 (2008).
45. Price, J. J. & Eaton, M. D. Reconstructing the evolution of sexual dichromatism: current color diversity does not reflect past rates of male and female change. *Evolution* **68**, 2026–2037 (2014).
46. Burns, K. J. A phylogenetic perspective on the evolution of sexual dichromatism in tanagers (Thraupidae): the role of female versus male plumage. *Evolution* **52**, 1219–1224 (1998).
47. Johnson, A. E., Jordan Price, J. & Pruett-Jones, S. Different modes of evolution in males and females generate dichromatism in fairy-wrens (Maluridae). *Ecol. Evol.* **3**, 3030–3046 (2013).
48. Dunn, P. O., Armenta, J. K. & Whittingham, L. A. Natural and sexual selection act on different axes of variation in avian plumage color. *Sci. Adv.* **1**, e1400155 (2015).
49. Rubenstein, D. R. Sexual and social competition: broadening perspectives by defining female roles. *Phil. Trans. R. Soc. B.* **367**, 2248–2252 (2012).
50. Alerstam, T., Hedenström, A. & Åkesson, S. Long-distance migration: evolution and determinants. *Oikos* **103**, 247–260 (2003).
51. Simpson, R. K., Johnson, M. A. & Murphy, T. G. Migration and the evolution of sexual dichromatism: evolutionary loss of female colouration with migration among wood-warblers. *Proc. R. Soc. B* **282**, 20150375 (2015).
52. Fitzpatrick, S. Colourful migratory birds: evidence for a mechanism other than parasite resistance for the maintenance of 'good genes' sexual selection. *Proc. R. Soc. Lond. B* **257**, 155–160 (1994).
53. Dunning, J. B. *CRC Handbook of Avian Body Masses* 2nd edn (CRC, 2008).
54. Valcu, M., Dale, J. & Kempenaers, B. rangeMapper: a platform for the study of macroecology of life-history traits. *Glob. Ecol. Biogeogr.* **21**, 945–951 (2012).
55. Land Processes Distributed Active Archive Center (LP DAAC). *MODIS/Terra Land Surface Temperature/Emissivity 8-Day L3 Global 0.05Deg* (LP DAAC, 2014).
56. Marchant, S. & Higgins, P. J. *Handbook of Australian, New Zealand & Antarctic Birds* (Oxford Univ. Press, 1990–2006).
57. Cramp, S. & Simmons, K. E. L. *Handbook of the Birds of Europe, the Middle East and North Africa: the Birds of the Western Palearctic* (Oxford Univ. Press, 1977–1994).
58. Brown, L. H., Urban, E. K. & Newmann, K. *The Birds of Africa* (Academic, 1982–2004).
59. Hockey, P. A. R., Dean, W. R. J. & Ryan, P. G. *Roberts — Birds of Southern Africa* 7th edn (John Voelcker Bird Book Fund, 2005).
60. Poole, A. & Gill, F. *Birds of North America* (Cornell Lab of Ornithology, 1992–2003).
61. Dunn, P. O., Whittingham, L. A. & Pitcher, T. E. Mating systems, sperm competition, and the evolution of sexual dimorphism in birds. *Evolution* **55**, 161–175 (2001).
62. Pitcher, T. E., Dunn, P. O. & Whittingham, L. A. Sperm competition and the evolution of testes size in birds. *J. Evol. Biol.* **18**, 557–567 (2005).
63. Cockburn, A. Prevalence of different modes of parental care in birds. *Proc. R. Soc. B* **273**, 1375–1383 (2006).
64. Martins, E. P. & Hansen, T. F. Phylogenies and the comparative method: a general approach to incorporating phylogenetic information into the analysis of interspecific data. *Am. Nat.* **149**, 646–667 (1997).
65. Freckleton, R. P., Harvey, P. H. & Pagel, M. Phylogenetic analysis and comparative data: a test and review of evidence. *Am. Nat.* **160**, 712–726 (2002).
66. Pagel, M. Inferring the historical patterns of biological evolution. *Nature* **401**, 877–884 (1999).
67. Hackett, S. J. *et al.* A phylogenomic study of birds reveals their evolutionary history. *Science* **320**, 1763–1768 (2008).
68. Garland, T. Jr & Ives, A. R. Using the past to predict the present: confidence intervals for regression equations in phylogenetic comparative methods. *Am. Nat.* **155**, 346–364 (2000).
69. Paradis, E., Claude, J. & Strimmer, K. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* 20, 289–290 (2004).
70. Pinheiro, J., Bates, D., DebRoy, S. & Sarkar, D. nlme: Linear and nonlinear mixed effects models. R package version 3.1–117. http://CRAN.R-project.org/package=nlme (2014).
71. Revell, L. J. Phylogenetic signal and linear regression on species data. *Methods Ecol. Evol.* **1**, 319–329 (2010).
72. Revell, L. J. Size-correction and principal components for interspecific comparative studies. *Evolution* **63**, 3258–3268 (2009).
73. Revell, L. J. phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecol. Evol.* **3**, 217–223 (2012).

74. R Core Team. *R: A language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2014).

75. Schielzeth, H. Simple means to improve the interpretability of regression coefficients. *Methods Ecol. Evol.* **1,** 103–113 (2010).

76. Bartoszek, K., Pienaar, J., Mostad, P., Andersson, S. & Hansen, T. F. A phylogenetic comparative method for studying multivariate adaptation. *J. Theor. Biol.* **314,** 204–215 (2012).

77. Beaulieu, J. M., Jhwueng, D. C., Boettiger, C. & O'Meara, B. C. Modeling stabilizing selection: expanding the Ornstein-Uhlenbeck model of adaptive evolution. *Evolution* **66,** 2369–2383 (2012).

78. Hansen, T. F. in *Modern Phylogenetic Comparative Methods and Their Application in Evolutionary Biology* (ed. Garamszegi, L. Z.) 351–379 (Springer, 2014).

79. Clavel, J., Escarguel, G. & Merceron, G. mvMORPH: an R package for fitting multivariate evolutionary models to morphometric data. *Meth. Ecol. Evol.* http://dx.doi.org/10.1111/2041-210X.12420 (2015).

80. Shipley, B. A new inferential test for path models based on directed acyclic graphs. *Struct. Equ. Modeling* **7,** 206–218 (2000).

81. Shipley, B. *Cause and Correlation in Biology: a User's Guide to Path Analysis, Structural Equations and Causal Inference* (Cambridge Univ. Press, 2002).

82. Gonzalez-Voyer, A. & von Hardenberg, A. in *Modern Phylogenetic Comparative Methods and Their Application in Evolutionary Biology* (ed. Garamszegi, L. Z.) 201–229 (Springer, 2014).

83. Shipley, B. The AIC model selection method applied to path analytic models compared using a d-separation test. *Ecology* **94,** 560–564 (2013).

84. Arnold, T. W. Uninformative parameters and model selection using Akaike's information criterion. *J. Wildl. Mgmt.* **74,** 1175–1178 (2010).

85. Burnham, K. P. & Anderson, D. R. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach* (Springer, 2002).
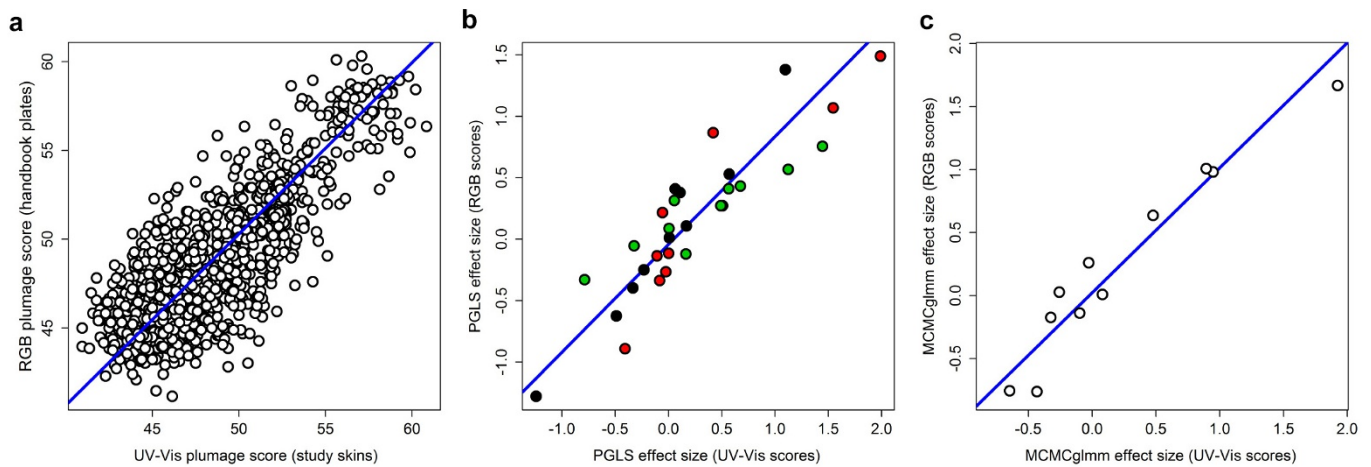
Extended Data Figure 1 | Comparison of plumage scores determined with handbook plates in RGB colour space with plumage scores determined with study skins using UV–Vis spectrometry. a, RGB versus UV–Vis scores calculated with 534 Australian bird species (reduced major axis (RMA) regression: $y = 0.965x + 2.025$, $N = 1068$, $R^2 = 0.670$, $P < 0.0001$). **b**, PGLS model effect sizes determined with RGB scores versus effect sizes determined with UV–Vis scores (RMA regression: $y = 0.878x - 0.042$, $N = 30$ model effects, $R^2 = 0.809$, $P < 0.0001$, each model had 305 Australian passerine species in it and black, red and green points reflect models predicting female scores, male scores and dichromatism scores, respectively). **c**, MCMCglmm model effect sizes (including interaction effects) determined with RGB versus UV–Vis scores (RMA regression: $y = 0.991x - 0.028$, $R^2 = 0.930$, $P < 0.0001$, $N = 11$ effects, the model had 305 Australian passerine species in it and each point reflects the mean effect size calculated from five models using a separate phylogenetic tree from http://birdtree.org each).
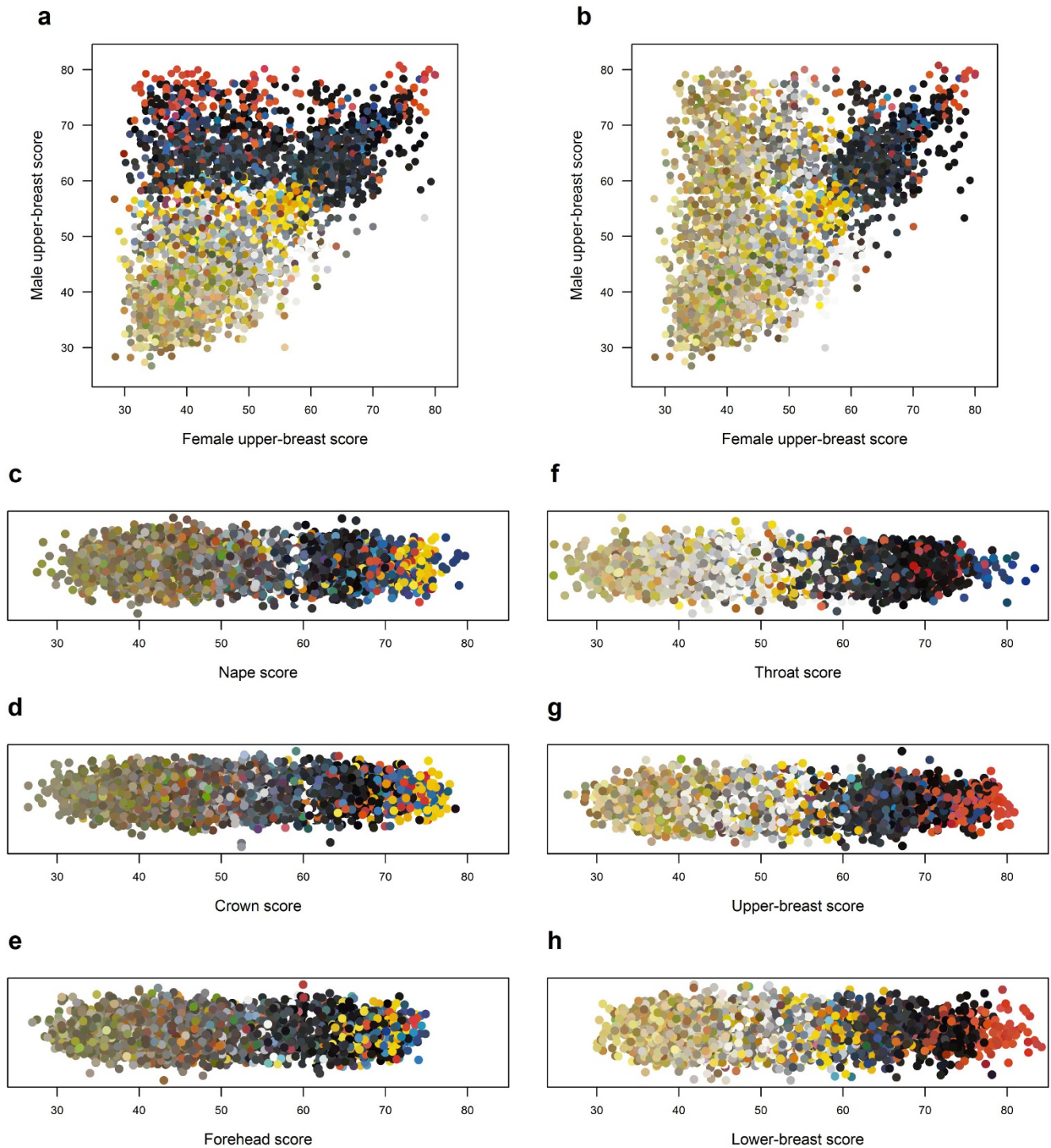
Extended Data Figure 2 | Basic patterns of plumage colouration in the order Passeriformes ($N = 5,983$ species). a, Male versus female patch scores (upper breast shown as an example). Data points are coloured by the male RGB values scored from handbook plates. **b**, As in **a**, only points are coloured with female RGB scores. **c–h**, The colours associated with different plumage scores differentiated by patch type. The y axis is a normally distributed random number used to spread out variation and improve visualization. The figure reflects patterns that agree with our intuition: dull greens, olives and browns have low plumage scores (are female-like), while richer or high contrasting colours (blacks, purples, blues, reds and yellows) have high scores (are male-like).
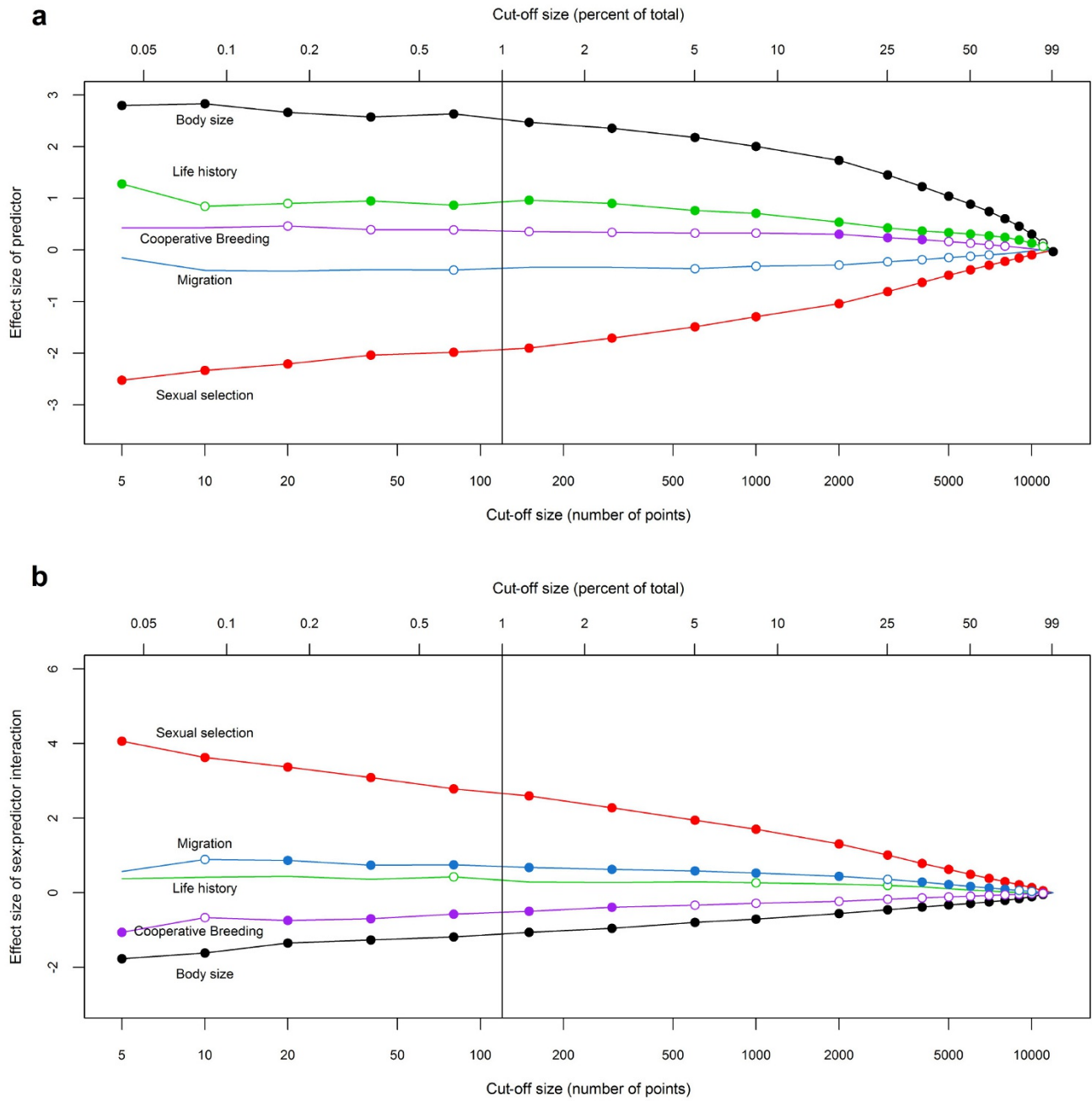
Extended Data Figure 3 | MCMCglmm results are robust to various cut-offs used to determine plumage scores. a, b, Main effects (**a**) and interaction (**b**) with sex effects. Filled circles: estimates where $P < 0.001$; open circles: estimates where $0.05 > P > 0.001$. The vertical line represents the cut-off value used in the main analysis. To reduce processing time for this analysis each MCMCglmm model ran for 10,000 iterations with a sampling interval of 100. Parameter estimates from the shortened runs were highly similar to the parameter estimates reported in the main MCMCglmm analysis (see Extended Data Table 4). Note that the observed decline in effect sizes as the cut-off size increases is the automatic outcome of how plumage scores are calculated. As the cut-off size approaches 100% the variance in plumage scores necessarily approaches 0 and so the effect sizes will inevitably also approach 0.
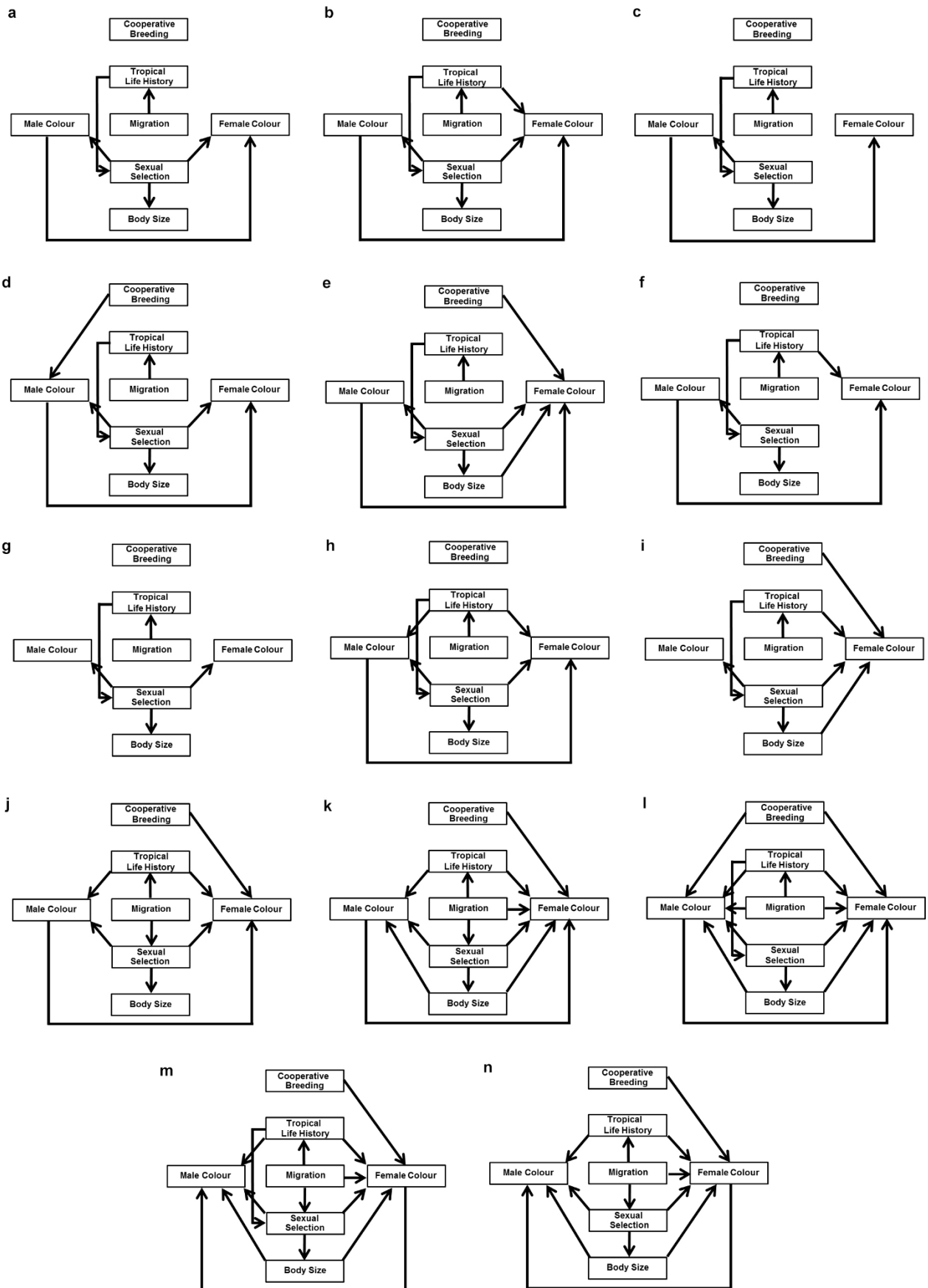
Extended Data Figure 4 | Candidate models for phylogenetic confirmatory path analysis. a–n, Arrows indicate hypothesized direct links between variables.

Extended Data Table 1 | Comparison of multivariate co-evolutionary models of male and female plumage ornamentation

a

Model	α	σ	log likelihood	AIC	ΔAIC
3	correlated	correlated	-39068	78153	-
2	independent	correlated	-39101	78217	64
1	independent	independent	-40140	80292	2139
5	-	correlated	-43481	86973	8820
4	-	independent	-45156	90321	12168

b

Parameter	Estimate
θ	female = 47.52 male = 51.02
α	$\begin{pmatrix} 2.30 & 0.90 \\ 0.90 & 1.22 \end{pmatrix}$
σ	$\begin{pmatrix} 290.55 & 216.50 \\ 216.50 & 222.17 \end{pmatrix}$

a, Multivariate Ornstein–Uhlenbeck (OU) models are ordered by Akaike information criterion (AIC) value from best (top) to worst. The α matrix determines the strength of selection and can incorporate an interaction among traits ('correlated') or can model independent selection on each trait ('independent'). The σ matrix determines the amount of variation (drift) in the stochastic evolutionary process, which can be independent or correlated among traits. Brownian motion models (4 and 5) do not contain α parameters. **b**, Parameter estimates for the best model (Model 3 in **a**). θ values are the estimated trait optima towards which male and female plumage colour evolve under OU processes. The α and σ matrices are shown for female-specific values in the top left cell, male-specific values in the bottom right cell and the interaction between trait values in the off-diagonal elements.

Extended Data Table 2 | Phylogenetic generalized least-squares models on predictors of sexual dichromatism in the Passeriformes

	Variable	N species	β	Std.Error	t-value	P
a, Single predictor models	Body Mass	5066	-0.76	0.17	-4.39	< 0.0001
	Wing Length	2855	-0.66	0.25	-2.66	0.011
	Degrees from equator	5831	0.60	0.09	6.77	< 0.0001
	Seasonality	5831	0.64	0.09	7.02	< 0.0001
	Clutch size	3909	0.51	0.13	4.01	< 0.0001
	Social polygyny	3504	1.51	0.15	9.83	< 0.0001
	Sexual size dimorphism	2855	0.70	0.15	4.59	< 0.0001
	Paternal care	5831	-0.99	0.12	-8.29	< 0.0001
	Cooperative Breeding	5831	-0.25	0.09	-2.77	0.007
	Migration	5830	0.81	0.09	9.37	< 0.0001
b, Multiple predictor model	Body size	2471	-0.95	0.25	-3.70	0.0002
	Tropical Life History	2471	-0.28	0.20	-1.40	0.16
	Sexual Selection	2471	1.76	0.19	9.20	<0.0001
	Cooperative Breeding	2471	-0.28	0.14	-2.01	0.04
	Migration	2471	0.81	0.19	4.33	<0.0001

To account for phylogenetic non-independence, models included a phylogenetic correlation structure using phylogenetic trees from http://birdtree.org[30]. Provided are the model-averaged predictions of 20 models (each using a different tree), with sexual dichromatism (that is, male plumage score − female plumage score) as the dependent variable. The phylogenetic signal, estimated as Pagel's λ coefficient, was $\lambda = 0.80$ (mean; range: 0.76–0.84). Predictor variables were scaled to a mean of 0 and standard deviation of 1.

Extended Data Table 3 | Phylogenetic principal component analysis loadings and the variance explained by each component

Predictor	Variable	Component Loadings		
		PPC1	PPC2	PPC3
Body size	Body mass	0.48	-0.13	
	Wing length	0.48	0.13	
	% Variance	93.15	6.85	
Tropical life history	Clutch size	-0.39	0.62	0
	Degrees from equator	-0.46	-0.27	-0.3
	Seasonality	-0.46	-0.26	0.3
	% Variance	76.59	17.32	6.09
Sexual selection	Sexual size dimorphism	-0.37	0.85	-0.01
	Social polygyny	-0.57	-0.27	0.51
	Paternal care	0.57	0.28	0.5
	% Variance	53.81	29.15	17.05

Provided are the mean values obtained from phylogenetic principal component analysis of 20 separate phylogenetic trees from http://birdtree.org. The first components were used in the main analysis. PPC, phylogenetic principal component.

Extended Data Table 4 | Morphological, life-history and social correlates of plumage colour scores in passerine birds ($N = 2{,}471$ species)

Predictor	Posterior mean	Lower 95% C.I.	Upper 95% C.I.	N	pMCMC
(Intercept)	45.91	40.64	51.33	987	< 0.001
sex	4.24	3.95	4.53	1018	< 0.001
body size	2.60	2.16	3.06	1023	< 0.001
life history	0.91	0.53	1.30	1024	< 0.001
sexual selection	-1.95	-2.30	-1.61	1009	< 0.001
cooperative breeding	0.40	0.13	0.68	1045	0.005
migration	-0.35	-0.72	0.01	1015	0.063
sex : body size	-1.14	-1.44	-0.86	1024	< 0.001
sex : life history	0.29	-0.09	0.67	1016	0.14
sex : sexual selection	2.70	2.41	3.00	1007	< 0.001
sex : cooperative breeding	-0.50	-0.80	-0.21	1018	0.0015
sex : migration	0.70	0.32	1.07	1015	0.0011

Shown are posterior estimates of the effect size in the full model (plus the 95% credibility intervals, the effective sample size (N), and the probability that the effect is different from the null hypothesis). The full model was the deviance information criterion (DIC) favoured model out of all possible additive models ($N = 243$ models) that include sex as well as all possible combinations of the five predictor variables and their respective interactions with sex. Values are means of 100 separate runs on each of 100 separate phylogenetic trees from http://birdtree.org.

Extended Data Table 5 | Comparison of models used in phylogenetic confirmatory path analysis

Model	C	ClCc	ΔClCc
K	28.81	67.12	0
L	27.90	68.25	1.13
N	43.04	81.35	14.23
M	42.13	82.47	15.35
J	169.00	201.23	134.11
E	184.74	214.93	147.81
H	189.26	215.41	148.29
B	207.37	235.54	168.42
D	244.17	274.37	207.25
A	277.68	303.83	236.71
F	304.60	330.75	263.63
C	377.06	401.19	334.07
I	790.13	820.32	753.20
G	1054.12	1078.24	1011.12

Models are ordered by ClCc value from best (top) to worst. *C* indicates Fisher's *C*-statistic.
For a visual illustration of the models see Extended Data Fig. 4.