SUPPLEMENTARY INFORMATION

Title: Global incidence of female birdsong is predicted by territoriality and biparental care in songbirds

Author list: Karan J. Odom[1,2,3]*, Marcelo Araya Salas[4,5,3], Lauryn Benedict[6], Kristi Lim[7], James Dale[8], Wesley H. Webb[8,9], Catherine Sheard[10], Joseph A. Tobias[11], Gregory F. Ball[2], Michelle L. Hall[12,13,14], Naomi E. Langmore[15], Michael S. Webster[3], Katharina Riebel[16]

Affiliations:

1. Department of Biological Sciences, University of the Pacific, Stockton, California, USA
2. Department of Psychology, University of Maryland, College Park, College Park, Maryland, USA
3. Cornell Lab of Ornithology and Department of Neurobiology and Behavior, Cornell University, Ithaca, New York, USA
4. Centro de Investigación en Neurociencias, Universidad de Costa Rica, San José, Costa Rica
5. Esciela de Biología, Universidad de Costa Rica, San José, Costa Rica
6. School of Biological Sciences, University of Northern Colorado, Greeley, Colorado, USA
7. College of Arts and Sciences, Cornell University, Ithaca, NY, USA
8. School of Environmental and Animal Sciences, Unitec, Auckland, New Zealand
9. School of Natural Sciences, Massey University, Auckland, New Zealand
10. School of Biological Sciences, University of Aberdeen, Aberdeen AB24 2TZ, United Kingdom
11. Department of Life Sciences, Silwood Park, Ascot, Berkshire, SL5 7PY, UK
12. School of BioSciences, The University of Melbourne, Melbourne, Vic., Australia
13. Bush Heritage Australia, Melbourne, Vic., Australia
14. School of Biological Sciences, The University of Western Australia, Perth, WA, Australia
15. Research School of Biology, Australian National University, Canberra, ACT, Australia
16. Institute of Biology, Leiden University, 2333 BE, Leiden, The Netherlands

Keywords: female, birdsong, evolution, songbirds, life-history

*Corresponding author:
Karan J. Odom
kodom@pacific.edu
Department of Biological Sciences
University of the Pacific
3601 Pacific Avenue
Stockton, CA, 95211

Abstract

Pronounced sexual dimorphism is generally assumed to evolve through sexual selection for elaborate male traits. However, there is increasing evidence that sexual dimorphism in traits such as birdsong may also evolve through loss of elaboration in females, but the evolutionary drivers underlying this process are obscure. Here we analyse ecological and natural history traits for over 1300 songbird species and show that increased female song incidence and elaboration are most directly associated with year-round territoriality, biparental care, and large body size. Phylogenetic path analysis indicates that mating system and breeding latitude primarily have indirect effects on female song evolution. Stable, tropical life histories and mating systems with biparental care promote female song, whereas evolutionary transitions to migration, reduced territoriality, and loss of male care led to losses or reductions of female song incidence. Our analyses provide a comprehensive framework for studying the drivers of sex differences and similarities in birdsong and reveal novel interactions among natural history and sexual selection pressures that have been hypothesized to independently shape elaborate traits.

Supplementary methods

Female song scores

The species included in the current study were compiled from previous studies on female song and species with evidence of female song from recording collections (Macaulay Library and xeno-canto; Garamszegi et al. 2007; Benedict, 2008; Price et al., 2009; Odom et al., 2014; Webb et al., 2016). We used Birds of the World and regional species accounts and field guides to score female song (Isler & Isler, 1987; Cramp et al., 1993; Jaramillo & Burke, 1999; Higgins et al., 2001, 2006; Hockey et al., 2005; Del Hoyo et al., 2010; Billerman et al., 2022). See the Methods in the main text for a description of the scoring approach.

Predictor variables

The predictor variables used to test hypotheses associated with female song were compiled from several sources including: (1) daily nest predation rates from Unzeta et al. 2020, (2) life- and natural history traits from Dale et al. 2015, and (3) territoriality and duet data from Tobias et al. 2016. A brief description of each predictor variable is below.

Daily nest predation rates – Daily nest predation rates were used to estimate, standardize, and compare predation threat for each species (Mayfield 1961). Data in the current study were predominantly compiled by Unzeta et al. 2020 and included calculations from both personal field data (as exposure days calculated by the Mayfield method; Mayfield 1961) and nest success reported from the literature. Cavity nesters were omitted from Unzeta et al. 2020, but added to the current study when available from field studies (T. Martin unpublished data) and the literature (Martin 1995, Martin and Clobert 1996, Remes et al. 2012).

Breeding latitude – Each species' geographical location was computed as the latitude (degrees from equator) of the breeding range centroid. This was calculated from polygons of breeding range maps transformed from an unprojected coordinate system (latitude, longitude) to an equal area projected coordinate system (cylindrical equal-area, latitude of the origin: 0°).

Body size (log mass) – Body mass data was collated from Dunning (2008). When more than one body mass entry was available for a species, the weighted mean was calculated per species by sample size. These values usually included both males and females, depending on availability. These data were log-transformed prior to statistical analysis.

Sexual size dimorphism – Sexual size dimorphism was calculated as the log (male wing length) − log(female wing length) to provide a proportional index of relative sizes of the sexes. Positive values reflect species where males are larger than females. See Dale et al. 2007 for a detailed explanation of this metric.

Biparental care – Biparental care was scored as 0 = absent or 1 = present primarily based on data provided in Cockburn 2006 and modified by Dale et al. 2015, including both the known and inferred data categories. For species in our data set not present in Cockburn 2006, additional parental care scores were obtained from del Hoyo et al. 2003-2011.

Cooperative breeding – Cooperative breeding was scored as 0 = absent or 1 = present also primarily based on data from Cockburn 2006 and modified by Dale et al. 2015. For species in our data set not present in Cockburn 2006, additional cooperative breeding scores were obtained from del Hoyo et al. 2003-2011.

Social mating system – Social polygyny was scored on a four-point scale following Owens and Hartley 1998, with 0 = strict social monogamy, 1 = monogamy with infrequent instances of polygyny (<5% of males; e.g., lazuli bunting, *Passerina amoena*), 2 = mostly social monogamy with regular facultative social polygyny (5 to 20% of males; e.g., American redstart, *Setophaga ruticilla*), and 3 = obligate resource defence polygyny (>20% of males; e.g., red-winged blackbird, *Agelaius phoeniceus*) or lek polygyny (e.g., Raggiana Bird-of-paradise, *Paradisaea raggiana*). Assignments were made based on species accounts (del Hoyo et al. 2003-2011, Marchant and Higgins 1990–2006, Cramp & Simmons 1977–1994, Brown et al. 1982–2004, Hockey et al. 2005, Poole & Gill 1992–2003) and previous comparative studies (Dale et al. 2007, Dunn et al. 2001, and Pitcher et al. 2005). A small number of passerine species with polygynandrous mating systems (e.g., the dunnock, *Prunella modularis* or sickle-billed vanga *Falculea palliata*) 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.

Migratory behaviour – Migration was scored on a scale from 0 to 2, with 0 = resident (breeding and non-breeding ranges identical), 1 = partial migration (some overlap between breeding and non-breeding ranges), 2 = complete (full) migration (no overlap between breeding and non-breeding ranges). Assignments were made based on the range maps within del Hoyo et al. 2003-2011.

Territoriality – Species were classified as 1 = non-territorial, 2 = seasonally or weakly territorial, or 3 = year-round territorial. We defined year-round territoriality as territory defence lasting throughout the year rather than residency within a restricted area, including migrants that are territorial on both the breeding and non-breeding grounds. Species that are vocal and aggressive (responsive to playbacks) for part of the year but remain in the same area silently for the rest of the year were classified as seasonal rather than year-round territorial. Seasonal or weak territoriality primarily included species with broadly overlapping home ranges or that joined mixed species flocks. Non-territorial species never defend territories and included species that defend a very small area around a nest site.

Duetting – Duets were scored by Tobias et al. (2016) as 0 = absent or 1 = present for each species and were defined as acoustic signals involving a male and female. In line with previous work, duets had to be composed of long-range acoustic signals that are coordinated or stereotyped in some way, whether they be loosely synchronous, regularly alternating, or precisely interwoven. While duets of songbirds are often comprised of songs, duets could include

other long-range vocalizations with song-like functions. Duet presence/absence was scored based on a variety of sources, including published literature, field observations, regional experts, and sound archives. Our duet score did not include species with male-male duets. For simplicity we removed species without female song in which females only contribute simple call-like notes from our female song / duet comparison. However, 20 of these 24 species with duets but no female song were monogamous, year-round territorial, and nonmigratory with biparental care, the same as most of other duetting species.

In addition to the above predictor variables, we measured seasonality (seasonal climate variation) and wing length (see Dale et al. 2015 for a description of these variables). These two variables were highly correlated ($r \geq 0.85$) with latitude and Log mass, respectively, and we therefore left them out of final analyses.

Statistical analysis

Treatment of highly correlated variables – We ran all statistical analyses using the individual predictor variables rather than using data reduction to concatenate correlated predictor variables (e.g., principal component analysis). We chose this approach because we wanted to directly evaluate associations among the predictor variables themselves and our measures of female song. We found this approach particularly valuable for path analysis, in which we assessed causal associations in line with our hypotheses. Prior to running statistical tests, we assessed correlations among predictor variables. If two predictor variables were highly correlated with a Pearson correlation coefficient above 0.7, we dropped one of the two highly correlated variables (Dormann et al. 2012). We kept variables that most closely aligned with the hypotheses for female song evolution. This resulted in dropping two predictor variables from further analyses: log wing length and seasonality (measured as the coefficient of variation of annual temperature). Wing length was highly correlated with body mass ($r = 0.93$) and seasonality was highly correlated with breeding latitude ($r = 0.85$; Fig. S11).

We also evaluated correlations among the three female song metrics to ensure that they were capturing statistically independent aspects of female singing behaviour. We removed species without female song during this assessment, as these values were the same for all three metrics and inflated their correlation. Correlations among the three response variables were low to intermediate ($r = 0.33$-0.58; Fig. S11), suggesting some non-independence, but also potential for each female song metric to have independent associations with the predictor variables.

Statistical models and assessment – We conducted the following analyses to compare each female song metric as the response variable to the following predictor variables: (1) phylogenetically informed Bayesian univariate regression models comparing daily nest predation rates, (2) phylogenetically informed Bayesian multivariate regression models with all remaining predictor variables, and (3) phylogenetically informed path analyses to evaluate causal pathways or associations among predictor variables. Daily nest predation rate had a much smaller sample size than the other parameters. Therefore, it was evaluated in its own model to preserve statistical power in analyses with the other predictor variables. All statistical analyses were conducted in R [71]. We built all our phylogenetically informed Bayesian regression models with brms (Bürkner 2017) and MCMCglmm (Hadfield 2010). Both packages allow phylogeny to be incorporated as a random effect, specifically as a species co-variance matrix into a mixed model Bayesian framework (Mai & Zhang 2018). Both analyses produced similar results (Tables 1, 2, S1 & S4), which were also supported by path analysis (Fig. 4). These two packages are known to produce nuanced differences in results (Mai and Zhang 2018). In these instances, we deferred to brms

results because we were able to specify a more appropriate model family for our ordinal response data and incorporate phylogenetic uncertainty in this package.

For analyses in brms, models were replicated on the 100 trees generated by birdtree.org using the R package brmsish (Araya-Salas 2022) to incorporate phylogenetic uncertainty and estimates were derived from the combined posterior distributions. Effect sizes are presented as median posterior estimates and 95% credibility intervals as the highest posterior density interval. Predictors in which the credible intervals of effect sizes did not include zero were regarded as having an effect on the response variable. We specified family in the following ways for the following response variable structures: ordinal response variables (incidence, elaboration and length) = cumulative (probit) distribution and presence/absence = bernoulli (logit) distribution. All prior structures were weakly informative for our brms models for both the population-level and group-level (phylogenetic) effects. We specified priors = normal(0, 2), class = "b". For each test, a preliminary model with four chains was run in order to determine the minimum number of iterations that will produce a reliable result. Performance was checked visually by plotting the trace and distribution of posterior estimates for all chains (supplementary material) and we produced posterior predictive check plots to evaluate model fit (Fig. S12). We also plotted the autocorrelation of successive sampled values to evaluate the independence of posterior samples. A potential scale reduction factor was used to assess model convergence and $\hat{R}$ was kept below 1.05 for all parameter estimates. We then used that number of iterations to run the final models over 100 trees (each with a single chain) to account for phylogenetic uncertainty. The effective sample size was kept above 3000 for all parameters.

For analyses in MCMCglmm, all models were run on a consensus tree incorporating all 100 trees. For our MCMCglmm models with more than one predictor variable, we performed model selection procedures using the function dredge in the R package MuMIn (Bartoń and Kamil 2019). Selection of the best model was determined based on DIC. If top models differed by less than 2 DIC values, we chose the model with the fewest parameters. We evaluated and report the final results for the top model. The appropriateness of using DIC for model selection is debated and discussed by Spiegelhalter et al. (2014), which is an additional reason that we accessed concordance with and place more emphasis on our brms results. For all MCMCglmm models we specified family = “categorical” and we used non-informative priors. We ran all models for 10,000 iterations, a burn-in of 600, and thinning of 10. MCMCglmm models will not run with missing predictor data. To compensate for missing predictor data, we centered and scaled all continuous predictor variables and set the missing values to zero, i.e., the mean. The main continuous predictor variable with missing data was sexual size dimorphism (n = 161). There was no missing data for other continuous predictor variables for species included in the final comparative analysis. If categorial predictor variable data was missing, that row was removed from the dataset. For most species in our study, there was no missing categorical predictor data. We visually examined model diagnostics for all models to ensure stationarity of trace plots and levels of autocorrelation <0.1 for all parameters. We used Raftery and Lewis’s diagnostic to determine the necessary number of iterations and Heidelberger and Welch’s diagnostic to ensure model convergence (Heidelberger and Welch 1983; Raftery and Lewis 1995). To assess final models, we used Bayesian p-values (pMCMC), a value analogous to traditional p-values. We accepted a cutoff of less than or equal to 0.05 to evaluate which variables contributed substantially to variation in each model.

Path analysis –We conducted phylogenetic path analysis using the R package PhyloPath (van der Bijl 2018). Phylogenetic confirmatory path analysis is a special case of structural

equation modelling, which allows causal relationships among multiple independent variables to be assessed in a phylogenetic framework (Shipley 2016, Gonzalez-Voyer & von Hardenberg 2014). This approach infers the most likely evolutionary pathways by considering both indirect and direct effects of predictor variables on response variables and the magnitude of these effects. Therefore, this technique has the ability to reveal the complex, interconnected ways that factors may interact in the natural world. PhyloPath only supports binary or continuous variables (not ordinal or multi-state categories). Therefore, we set our ordinal or categorical variables as numeric in R for our PhyloPath analyses to be able to use all variables and character states. Using PhyloPath, we simultaneously evaluated both the tropical natural history and sexual selection hypotheses and the interaction of predictor variables within each hypothesis (e.g., Fig. 1, Table S3). For each of the four predictor variables (incidence, elaboration, length and presence/absence), we compared 115 models within each analysis, representing the prominent combinations of connections between the predictor variables (see Fig. S13 for the major alternative models tested). Models were evaluated under the default Pagel's lambda model of evolution. We assessed model fit by comparing their C-statistic information criterion (CICc), which takes into account the number of parameters/paths (Gonzalez-Voyer and von Hardenberg 2014). All path analyses resulted in a single best model: this model had the lowest CICc and differed from all other models by more than 2 CICc values.

Phylogenetic signal – To directly evaluate the extent to which phylogenetic relationships explain expression of the female song metrics, we computed phylogenetic signal as Blomberg's K for each female song metric using the phylosig function in the phytools R package (Blomberg et al. 2003; Revell 2012). Because Blomberg's K requires continuous variables, we temporarily set our response variables as numeric for this analysis.

Comparison of female song to duets - To investigate how female song incidence is related to duet occurrence, we conducted two analyses: (1) we looked at whether female song incidence is associated with the occurrence of duetting, and (2) we examined which predictor variables are associated with female songs and duets. For this second analysis, we created an ordinal song variable that contained three levels: species without female song (female song absent), species with female song but that do not duet (female solo song), and species with both female song and duets (duetting species). Similar to the other regression analyses, we ran phylogenetically informed Bayesian regression models in brms and MCMCglmm. We used the same parameters, priors, and assessment as described for the other regression analyses above.

Interactions among predictor variables - We also evaluated potential interactions among key predictor variables to assess whether there are interactions among the hypotheses. In addition to the above path and regression analyses, we conducted follow up regression analyses including interactions between mating system and territoriality, as well as mating system and biparental care. This allowed us to assess if and in what way natural history and sexual selection theory interact to create the broad patterns of sexual dimorphism in song observed.

References for supplementary materials

Araya-Salas (2022), *brmsish: random stuff on brms bayesian models*. R package version 1.0.0.

Bartoń, B., & Kamil, M. (2019). Package "MuMIn.".

Benedict, L. (2008). Occurrence and life history correlates of vocal duetting in North American passerines. Journal of Avian Biology, 39(1), 57–65. https://doi.org/10.1111/j.0908-8857.2008.04103.x

Billerman, S. M., Keeney, B. K., Rodewald, P. G., & Schulenberg, T. S. (Eds.). (2022). Birds of the World. Cornell Laboratory of Ornithology. https://birdsoftheworld.org/bow/home

Blomberg, S. P., Garland Jr, T., & Ives, A. R. (2003). Testing for phylogenetic signal in comparative data: behavioral traits are more labile. *Evolution*, *57*, 717-745.

Brown, L. H., Urban, E. K. & Newmann, K. The Birds of Africa (Academic, 1982–2004).

Bürkner, P.-C. (2017). Brms: an R package for Bayesian multilevel models using Stan. J. Stat. Softw., 80, 1–28.

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).

Cockburn, A. Prevalence of different modes of parental care in birds. Proc. R. Soc. B 273, 1375–1383 (2006).

Dale, J. et al. Sexual selection explains Rensch's rule of allometry for sexual size dimorphism. Proc. R. Soc. B 274, 2971–2979 (2007).

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*, 367-370.

del Hoyo, J., Elliott, A. & Christie, D. A. Handbook of the Birds of the World Vols 8–16. (Lynx Edicions, 2003–2011).

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).
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).

Dunning, J. B. CRC Handbook of Avian Body Masses. 2nd edn (CRC Press, 2007).

Dormann, C.F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G. et al. 2012. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography 36: 27–46.

Garamszegi, L. Z., Pavlova, D. Z., Eens, M., & Møller, A. P. (2007). The evolution of song in female birds in Europe. Behavioral Ecology, 18(1), 86–96. https://doi.org/10.1093/beheco/arl047

Gonzalez-Voyer, A., & von Hardenberg, A. (2014). An introduction to phylogenetic path analysis. Modern phylogenetic comparative methods and their application in evolutionary biology: Concepts and practice, 201-229.

Hackett, S. J. et al. A phylogenomic study of birds reveals their evolutionary history. Science 320, 1763–1768 (2008).

Hadfield, J.D. (2010). MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. J. Stat. Softw., 33, 1–22.

Heidelberger P and Welch PD. Simulation run length control in the presence of an initial transient. Opns Res., 31, 1109-44 (1983)

Higgins, P. J., Peter, J. M., & Cowling, S. J. (2006). Handbook of Australian, New Zealand and Antarctic birds. In Boatbill to starlings (Vol. 7). Oxford University Press.

Higgins, P. J., Peter, J. M., & Steele, W. K. (2001). Handbook of Australian, New Zealand and Antarctic birds. In Tyrant-flycatchers to Chats (Vol. 5). Oxford University Press.

Hockey, P. A. R., Dean, W. R. J. & Ryan, P. G. Roberts — Birds of Southern Africa. 7th edn (John Voelcker Bird Book Fund, 2005).

Isler, M. I., & Isler, P. R. (1987). The Tanagers: natural history, distribution, and identification. Smithsonian Institution Press.

Jaramillo, A., & Burke, P. (1999). New World Blackbirds: The Icterids. Princeton University Press.

Mai, Y. & Zhang, Z. (2018). Software packages for Bayesian multilevel modeling. Struct. Equ. Model., 25, 650–658.

Marchant, S. & Higgins, P. J. Handbook of Australian, New Zealand & Antarctic Birds (Oxford Univ. Press, 1990–2006).

Martin, T. E. 1995. Avian life history evolution in relation to nest sites, nest predation, and food. Ecological Monographs 65:101–127.

Martin, T. E., and J. Clobert. 1996. Nest predation and avian life- history evolution in Europe versus North America: a possible role of humans? American Naturalist 147:1028–1046.

Mayfield, H. 1961. Nesting success calculated from exposure. Wilson Bulletin 73:255–261.
Odom, K. J., Hall, M. L., Riebel, K., Omland, K. E., & Langmore, N. E. (2014). Female song is widespread and ancestral in songbirds. Nature Communications, 5(1), 3379. https://doi.org/10.1038/ncomms4379

Owens, I. P., & Hartley, I. R. (1998). Sexual dimorphism in birds: why are there so many different forms of dimorphism?. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *265*, 397-407.

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).

Poole, A. & Gill, F. Birds of North America (Cornell Lab of Ornithology, 1992–2003)

Price, J. J., Lanyon, S. M., & Omland, K. E. (2009). Losses of female song with changes from tropical to temperate breeding in the New World blackbirds. Proceedings of the Royal Society B: Biological Sciences, 276, 1971–1980. https://doi.org/10.1098/rspb.2008.1626

R Core Team. (2019). *R: a language and environment for statistical computing*.

Raftery, A.E. and Lewis, S.M. (1995). The number of iterations, convergence diagnostics and generic Metropolis algorithms. In Practical Markov Chain Monte Carlo (W.R. Gilks, D.J. Spiegelhalter and S. Richardson, eds.). London, U.K.: Chapman and Hall.

Revell, L. J. (2012). phytools: an R package for phylogenetic comparative biology (and other things). Methods in Ecology and Evolution, (2), 217-223.

Shipley, B. (2016). Cause and correlation in biology: a user's guide to path analysis, structural equations and causal inference with R. Cambridge University Press.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & van der Linde, A. (2014). The deviance information criterion: 12 years on. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 76, 485–493. https://doi.org/10.1111/rssb.12062

Tobias, J. A., Sheard, C., Seddon, N., Meade, A., Cotton, A. J., & Nakagawa, S. (2016). Territoriality, social bonds, and the evolution of communal signaling in birds. *Frontiers in Ecology and Evolution*, *4*, 74.

Unzeta, M., Martin, T. E., & Sol, D. (2020). Daily nest predation rates decrease with body size in passerine birds. *The American Naturalist*, *196*, 743-754.

Webb, W. H., Brunton, D. H., Aguirre, J. D., Thomas, D. B., Valcu, M., Dale, J., & Webb, W. H. (2016). female song occurs in songbirds with more elaborate female coloration and reduced sexual dichromatism. 4 (March), 1–8. https://doi.org/10.3389/fevo.2016.00022

SUPPLEMENTARY TABLES & FIGURES

Response variable	Best model / Predictor variables	Posterior mean	Lower 95% CI	Upper 95% CI	Effective sample size	Bayesian p-value	
Incidence	log mass + latitude + Biparental care + cooperation + migration + territoriality						
	Rare: log mass	-0.14	-0.52	0.25	72	0.5072	
	Occasional: log mass	0.20	-0.05	0.46	203	0.1266	
	Regular: log mass	0.43	0.15	0.71	231	0.0026	**
	Same as male: log mass	0.93	0.65	1.24	244	<0.0001	***
	Rare: latitude	0.90	0.37	1.37	57	<0.0001	***
	Occasional: latitude	0.80	0.48	1.15	198	<0.0001	***
	Regular: latitude	0.46	0.11	0.80	248	0.0117	*
	Same as male: latitude	-0.45	-0.83	-0.09	245	0.0133	*
	Rare: Biparental care (yes)	0.94	-1.32	3.47	23	0.4410	
	Occasional: Biparental care (yes)	1.39	0.04	2.87	120	0.0326	*
	Regular: Biparental care (yes)	1.88	0.07	3.80	61	0.0290	*
	Same as male: Biparental care (yes)	4.33	1.92	7.04	30	<0.0001	***
	Rare: cooperation (yes)	-1.22	-2.46	-0.10	50	0.0157	*
	Occasional: cooperation (yes)	0.26	-0.34	0.83	249	0.3946	
	Regular: cooperation (yes)	0.54	-0.09	1.18	290	0.0871	.
	Same as male: cooperation (yes)	0.68	0.02	1.38	344	0.0354	*
	Rare: migration (partial)	1.15	0.16	2.13	85	0.0260	*
	Occasional: migration (partial)	0.43	-0.23	1.14	264	0.2143	
	Regular: migration (partial)	-0.77	-1.67	0.12	205	0.0891	.
	Same as male: migration (partial)	0.11	-0.80	1.05	225	0.8398	
	Rare: migration (full)	0.58	-0.39	1.69	83	0.3052	
	Occasional: migration (full)	-0.21	-1.01	0.58	246	0.6022	
	Regular: migration (full)	-0.87	-1.81	0.12	231	0.0791	.
	Same as male: migration (full)	-4.14	-10.97	-0.63	9	0.0129	*
	Rare: territoriality (seasonal)	0.66	-0.84	2.57	36	0.4026	
	Occasional: territoriality (seasonal)	-0.14	-1.04	0.71	225	0.7759	
	Regular: territoriality (seasonal)	0.55	-0.59	1.86	106	0.3372	
	Same as male: territoriality (seasonal)	0.51	-0.83	1.88	121	0.4567	
	Rare: territoriality (year-round)	2.43	0.55	4.43	29	0.0030	**
	Occasional: territoriality (year-round)	2.27	1.26	3.22	222	<0.0001	***
	Regular: territoriality (year-round)	3.41	2.14	4.68	129	<0.0001	***
	Same as male: territoriality (year-round)	4.10	2.69	5.49	153	<0.0001	***
Elaboration	log mass + latitude + mating system						
	A lot less: log mass	-0.14	-0.75	0.37	52	0.6610	
	Somewhat less: log mass	0.50	0.30	0.70	532	<0.0001	***
	Same as male: log mass	0.73	0.52	0.95	500	<0.0001	***
	A lot less: mating sys (<5% polygyny)	0.20	-1.69	2.17	77	0.7540	
	Somewhat less: mating sys (<5% polygyny)	0.10	-0.66	0.90	724	0.8120	
	Same as male: mating sys (<5% polygyny)	-1.10	-2.32	0.01	315	0.0540	.
	A lot less: mating sys (5-20% polygyny)	-0.94	-3.33	1.20	39	0.4730	
	Somewhat less: mating sys (5-20% polygyny)	0.22	-0.50	0.93	788	0.5420	

	Same as male: mating sys (5-20% polygyny)	-1.16	-2.24	-0.15	303	0.0220	*
	A lot less: mating sys (>20% polygyny)	-6.90	-18.34	-0.74	7	<0.0001	***
	Somewhat less: mating sys (>20% polygyny)	-2.89	-3.97	-1.80	237	<0.0001	***
	Same as male: mating sys (>20% polygyny)	-3.80	-4.91	-2.77	272	<0.0001	***
	A lot less: latitude	0.27	-0.29	0.93	38	0.4200	
	Somewhat less: latitude	0.14	-0.05	0.33	545	0.1770	
	Same as male: latitude	-0.27	-0.47	-0.07	591	0.0100	*
Song length	log mass + latitude + Biparental care + cooperation + migration + territoriality						
	A lot less: log mass	0.19	-0.27	0.66	53	0.4239	
	Somewhat less: log mass	0.33	0.06	0.60	184	0.0129	*
	Same as male: log mass	0.56	0.32	0.79	481	<0.0001	***
	More than male: log mass	1.32	0.53	1.95	20	0.0028	**
	A lot less: latitude	0.80	0.28	1.31	83	0.0044	**
	Somewhat less: latitude	0.04	-0.31	0.37	207	0.8060	
	Same as male: latitude	0.08	-0.22	0.37	594	0.6171	
	More than male: latitude	-0.45	-1.24	0.19	39	0.1823	
	A lot less: Biparental care (yes)	0.93	-1.65	3.52	34	0.4660	
	Somewhat less: Biparental care (yes)	2.39	0.52	4.78	34	0.0032	**
	Same as male: Biparental care (yes)	2.52	1.16	3.99	189	0.0004	***
	More than male: Biparental care (yes)	4.38	0.72	8.94	13	0.0022	**
	A lot less: cooperation (yes)	-0.74	-2.03	0.43	75	0.2006	
	Somewhat less: cooperation (yes)	-0.14	-0.87	0.52	179	0.7143	
	Same as male: cooperation (yes)	0.62	0.06	1.18	598	0.0314	*
	More than male: cooperation (yes)	-1.39	-3.27	0.33	27	0.0877	.
	A lot less: migration (partial)	-1.26	-3.09	0.33	43	0.1278	
	Somewhat less: migration (partial)	0.93	0.20	1.73	208	0.0143	*
	Same as male: migration (partial)	0.22	-0.44	0.90	603	0.5245	
	More than male: migration (partial)	1.80	-1.75	4.54	11	0.2757	
	A lot less: migration (full)	-0.59	-2.01	0.87	57	0.4670	
	Somewhat less: migration (full)	0.93	0.08	1.77	218	0.0316	*
	Same as male: migration (full)	-0.79	-1.75	0.16	286	0.1010	
	More than male: migration (full)	-2.00	-6.62	1.29	7	0.3581	
	A lot less: territoriality (seasonal)	1.01	-1.24	4.36	14	0.5205	
	Somewhat less: territoriality (seasonal)	0.35	-0.65	1.38	109	0.5187	
	Same as male: territoriality (seasonal)	-0.11	-0.86	0.67	497	0.7698	
	More than male: territoriality (seasonal)	-0.04	-2.19	2.01	26	0.8944	
	A lot less: territoriality (year-round)	4.32	1.98	7.33	17	<0.0001	***
	Somewhat less: territoriality (year-round)	3.19	2.14	4.29	121	<0.0001	***
	Same as male: territoriality (year-round)	2.44	1.59	3.30	485	<0.0001	***
	More than male: territoriality (year-round)	4.48	2.22	6.76	30	<0.0001	***

Table S1. MCMCglmm best model results for three metrics of female birdsong (incidence, elaboration and length) compared to natural history predictor variables. The association (or "significance") of a predictor variable with female song was assessed based on the following Bayesian p-values: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, . $p < 0.1$.

Predictor variable	Posterior mean	Lower 95% CI	Upper 95% CI	$\hat{R}$	Bulk_ESS
Daily nest predation rate (univariate model)	-0.274	-1.570	0.761	1.006	37960
Log mass	1.127	0.440	2.142	1.054	1015
Sexual size dimorphism	-0.450	-1.081	0.055	1.018	2945
Mating system (<5% polygyny)	-0.765	-2.240	0.577	1.015	3621
Mating system (5-20% polygyny)	-1.023	-2.412	0.180	1.017	3158
Mating system (>20% polygyny)	-0.047	-1.238	1.151	1.006	10451
Territoriality (seasonal)	3.481	1.996	5.647	1.045	1227
Territoriality (year-round)	0.935	0.127	1.991	1.027	2037
Biparental care (yes)	0.839	-0.704	2.477	1.030	1785
Cooperative breeding (yes)	0.442	-0.218	1.262	1.016	3508
Migration (partial)	0.128	-0.893	1.190	1.007	7906
Migration (full)	-0.878	-1.793	-0.159	1.019	2777
Breeding latitude	0.260	-0.294	0.889	1.014	3997

Table S2. Phylogenetic mixed model results processed in brms of female song present vs absent and natural history predictor variables. Predictor variables that are associated with whether female song is present / absent are **in bold** (these are variables whose posterior distributions, including 95% confidence intervals that do not cross zero).

Predictor variable	Estimate	l-95% CI	u-95% CI	$\hat{R}$	Bulk_ESS	Tail_ESS
Log mass	0.526	0.274	0.819	1.037	1470.495	3600.12
Sexual size dimorphism	-0.119	-0.34	0.09	1.02	2734.828	10667.49
Mating system (<5% polygyny)	0.354	-1.031	1.724	1.002	77181.51	682991.5
Mating system (5-20% polygyny)	-1.473	-2.751	-0.275	1.007	8475.02	26880.25
Mating system (>20% polygyny)	-0.983	-2.39	0.374	1.005	13652.09	265728.7
Territoriality (seasonal)	0.582	-0.109	1.327	1.01	5991.647	19017.05
Territoriality (year-round)	2.402	1.562	3.397	1.031	1781.771	4370.974
Biparental care (yes)	1.223	0.085	2.424	1.017	3299.364	11760.59
Cooperative breeding (yes)	0.126	-0.236	0.49	1.009	5935.942	17725.77
Migration (partial)	0.554	0.088	1.057	1.012	4675.982	12034.67
Migration (full)	0.234	-0.35	0.839	1.008	7758.959	33926.36
Breeding latitude	-0.099	-0.326	0.12	1.014	3870.436	11898.9
Mating system (<5% polygyny):Territoriality (seasonal)	-0.169	-1.621	1.304	1.002	58240.44	647609.1
Mating system (5-20% polygyny):Territoriality (seasonal)	1.758	0.432	3.172	1.01	5507.005	16215.36
Mating system (>20% polygyny):Territoriality (seasonal)	0.449	-1.068	1.965	1.004	16563.19	510245
Mating system (<5% polygyny):Territoriality (year-round)	-1.151	-2.973	0.649	1.002	61317.9	662893.5
Mating system (5-20% polygyny):Territoriality (year-round)	1.04	-0.885	3.018	1.002	48296.7	636060.9
Mating system (>20% polygyny):Territoriality (year-round)	-0.261	-3.331	2.79	1.001	1047104	669978.8

Table S3. Phylogenetic mixed model results processed in brms of female song incidence and natural history predictor variables, including an interaction between mating system and territoriality. Predictor variables that are associated with each of three female song metric in bold (these are variables whose posterior distributions, including 95% confidence intervals that do not cross zero).

Response variable	Level (Category)	Predictor variable	Posterior mean	Lower 95% CI	Upper 95% CI	Effective sample size	Bayesian p-value	
Incidence	Rare	Daily nest predation rate	-0.29	-1.02	0.39	122	0.4286	
	Occasional	Daily nest predation rate	-0.40	-0.92	0.13	202	0.1410	
	Regular	Daily nest predation rate	-0.22	-0.79	0.37	163	0.4787	
	Same as male	Daily nest predation rate	0.19	-0.30	0.69	263	0.4252	
Elaboration	A lot less	Daily nest predation rate	0.18	-0.55	0.88	175	0.6100	
	Somewhat less	Daily nest predation rate	-0.24	-0.66	0.15	507	0.2300	
	Same as male	Daily nest predation rate	-0.01	-0.43	0.41	537	0.9900	
Song length	A lot less	Daily nest predation rate	-3.14	-6.23	-0.22	10	0.0264	*
	Somewhat less	Daily nest predation rate	-0.07	-0.55	0.39	503	0.7787	
	Same as male	Daily nest predation rate	-0.47	-1.10	0.07	323	0.0857	.

Table S4. MCMCglmm Bayesian phylogenetic model results for daily nest predation rate and three female song metrics. The association (or "significance") of a predictor variable with female song was assessed based on the following Bayesian p-values: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, . $p < 0.1$.

Response variable	Predictor variable	Posterior mean	Lower 95% CI	Upper 95% CI	Effective sample size	Bayesian p-value	
Incidence	Rare: Duet present	-0.002	-1.59	1.10	24	0.89	
	Occasional: Duet present	1.58	0.92	2.20	83	<0.0001	***
	Regular: Duet present	3.51	2.92	4.11	162	<0.0001	***
	Same as male: Duet present	4.53	3.92	5.09	248	<0.0001	***
Elaboration	A lot less: Duet present	2.83	1.93	3.90	62	<0.0001	***
	Somewhat less: Duet present	2.61	2.09	3.16	334	<0.0001	***
	Same as male: Duet present	3.69	3.19	4.20	478	<0.0001	***
Song length	A lot less: Duet present	3.33	2.58	4.17	95	<0.0001	***
	Somewhat less: Duet present	3.12	2.57	3.66	238	<0.0001	***
		3.31	2.80	3.83	323	<0.0001	***
	Same as male: Duet present	4.64	3.40	6.03	29	<0.0001	***

Table S5. MCMCglmm Bayesian phylogenetic model results for duet presence/absence and three female song metrics. The association (or "significance") of a predictor variable with female song was assessed based on the following Bayesian p-values: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, . $p < 0.1$.

Full model / Predictor variables	Posterior mean	Lower 95% CI	Upper 95% CI	Effective sample size	Bayesian p-value	
No female song vs Female solo song vs Duet ~ log mass + SSD + latitude + mating system + Biparental care + cooperation + migration + territoriality						
Female solo only: log mass	0.40	0.20	0.60	50	<0.001	**
Duet: log mass	0.49	0.23	0.70	69	<0.001	**
Female solo only: sexual size dimorphism	-0.08	-0.30	0.09	66	0.4574	
Duet: sexual size dimorphism	0.02	-0.30	0.37	37	0.9085	
Female solo only: mating system (<5% polygyny)	0.20	-0.50	0.98	95	0.6170	
Duet: mating system (<5% polygyny)	-2.07	-4.70	0.19	11	0.0426	*
Female solo only: mating system (5-20% polygyny)	0.35	-0.37	1.03	82	0.3064	
Duet: mating system (5-20% polygyny)	-2.38	-6.82	1.00	3	0.3255	
Female solo only: mating system (>20% polygyny)	-0.66	-1.65	0.33	97	0.1745	
Duet: mating system (>20% polygyny)	-3.81	-6.24	-1.01	17	<0.001	**
Female solo only: Biparental care (yes)	0.66	-0.45	1.80	85	0.2277	
Duet: Biparental care (yes)	0.61	-1.69	2.95	23	0.7362	
Female solo only: cooperation (yes)	0.49	-0.01	0.96	57	0.0383	*
Duet: cooperation (yes)	-0.05	-0.74	0.50	72	0.9043	
Female solo only: migration (partial)	0.22	-0.25	0.67	124	0.3574	
Duet: migration (partial)	-0.41	-1.53	0.39	38	0.3936	
Female solo only: migration (full)	-0.30	-0.86	0.28	105	0.3043	
Duet: migration (full)	-1.67	-3.79	0.35	11	0.1213	
Female solo only: territoriality (seasonal)	0.13	-0.45	0.63	93	0.6468	
Duet: territoriality (seasonal)	1.72	0.03	3.61	16	0.0404	*
Female solo only: territoriality (year-round)	1.63	0.97	2.27	65	<0.001	**
Duet: territoriality (year-round)	6.11	4.50	8.13	18	<0.001	**
Female solo only: latitude	0.44	0.18	0.70	49	<0.001	**
Duet: latitude	-0.47	-0.81	-0.07	50	0.0043	**

Table S6. MCMCglmm Bayesian phylogenetic model results for female song and duets (scored ordinally) and natural history predictor variables. The association (or "significance") of a predictor variable with female song was assessed based on the following Bayesian p-values: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, . $p < 0.1$

Response variable	Model#	k	q	C	CICc	delta_CICc	l	w
Incidence	m111	20	25	123	174	0	1.00E+00	8.72E-01
	m112	21	24	129	178	3.94	1.40E-01	1.22E-01
	m109	22	23	137	184	9.85	7.26E-03	6.33E-03
	m106	21	24	148	197	23.11	9.58E-06	8.36E-06
	m105	22	23	151	198	24.01	6.10E-06	5.32E-06
	m102	22	23	154	201	27.05	1.33E-06	1.16E-06
	m101	23	22	162	207	32.97	6.92E-08	6.04E-08
	m108	22	23	168	216	41.35	1.05E-09	9.14E-10
	m110	22	23	170	217	43.11	4.36E-10	3.80E-10
	m107	23	22	179	224	49.36	1.91E-11	1.67E-11
Elaboration	m111	20	25	133	185	0	1.00E+00	9.34E-01
	m112	21	24	141	190	5.34	6.91E-02	6.46E-02
	m109	22	23	151	198	13.26	1.32E-03	1.24E-03
	m106	21	24	157	207	22.07	1.61E-05	1.51E-05
	m105	22	23	161	208	23.09	9.70E-06	9.06E-06
	m102	22	23	165	212	27.42	1.11E-06	1.04E-06
	m101	23	22	175	220	35.34	2.12E-08	1.98E-08
	m108	22	23	185	232	47.26	5.47E-11	5.11E-11
	m110	22	23	186	233	48.11	3.58E-11	3.34E-11
	m107	23	22	197	242	57.5	3.26E-13	3.05E-13
Length	m111	20	25	125	177	0	1.00E+00	9.30E-01
	m112	21	24	132	182	5.19	7.45E-02	6.93E-02
	m109	22	23	146	193	16.75	2.30E-04	2.14E-04
	m106	21	24	151	200	23.62	7.44E-06	6.92E-06
	m105	22	23	157	204	27.78	9.27E-07	8.63E-07
	m102	22	23	158	205	28.82	5.53E-07	5.15E-07
	m101	23	22	172	217	40.38	1.70E-09	1.58E-09
	m108	22	23	174	221	44.88	1.80E-10	1.67E-10
	m110	22	23	178	225	48.44	3.03E-11	2.82E-11
	m107	23	22	190	236	59.2	1.40E-13	1.30E-13
Present/Absent	m111	20	25	177	228	0	1.00E+00	8.67E-01
	m112	21	24	183	232	3.79	1.50E-01	1.30E-01
	m109	22	23	193	240	11.39	3.37E-03	2.92E-03
	m106	21	24	199	249	20.37	3.77E-05	3.27E-05
	m105	22	23	203	250	21.85	1.80E-05	1.56E-05
	m102	22	23	205	252	24.17	5.66E-06	4.90E-06

m101	23	22	215	260	31.76	1.27E-07	1.10E-07
m108	22	23	230	277	48.69	2.67E-11	2.32E-11
m110	22	23	231	278	49.43	1.85E-11	1.60E-11
m107	23	22	241	286	57.7	2.95E-13	2.56E-13
m104	22	23	248	295	66.67	3.34E-15	2.89E-15
m100	23	22	252	297	69.07	1.00E-15	8.71E-16

Table S7. PhyloPath best model CICc scores for the analyses of all four sets of analyses.

Fig. S1. Female song (a) elaboration and (b) length overlaid on a phylogenetic tree of all songbirds (passeri) for which we have sex-specific vocal information. Each legend shows the degree of female song behavior, including species with female song but no elaboration or length data in grey. Only terminal branch colors are accurate; internal branch colors are not an ancestral state reconstruction. We also show the number and total percent of species with each trait in the corresponding bar charts.

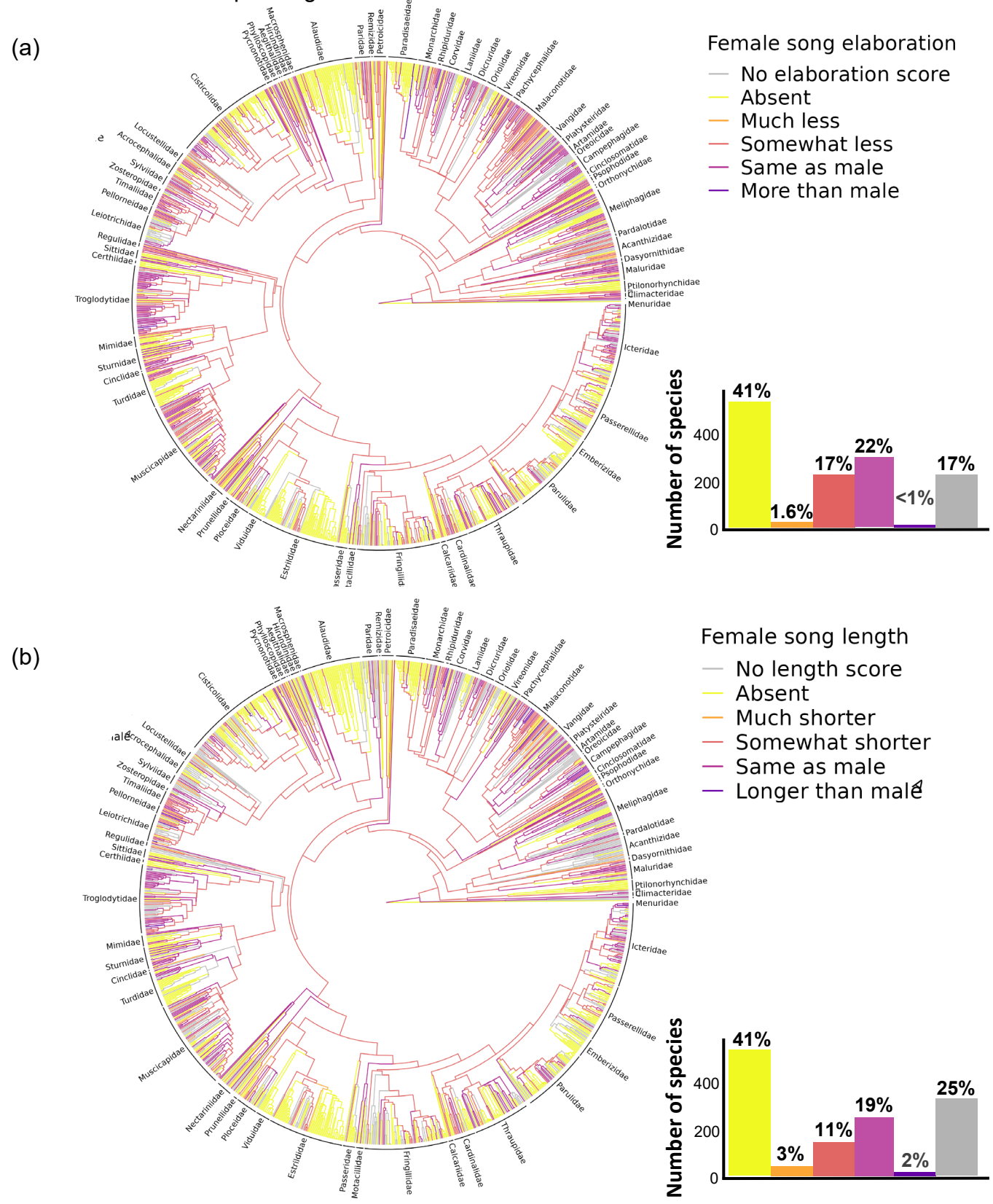
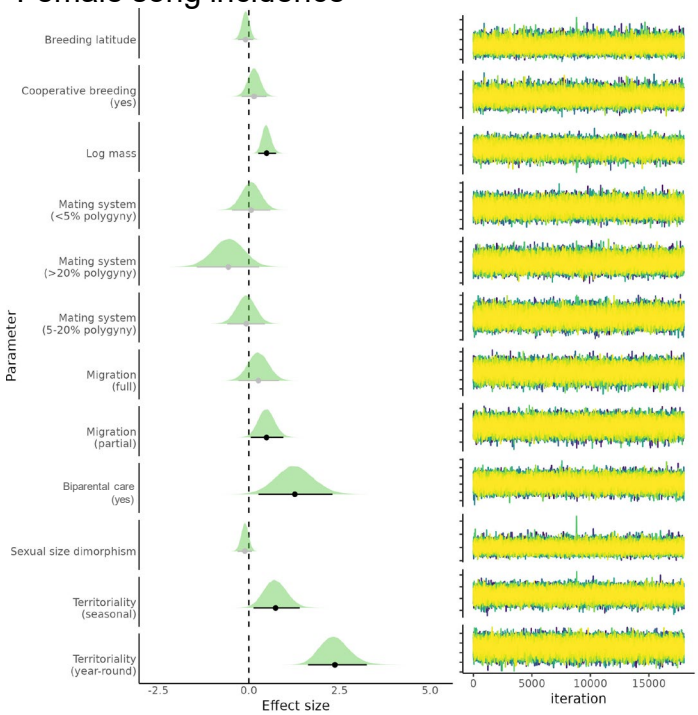
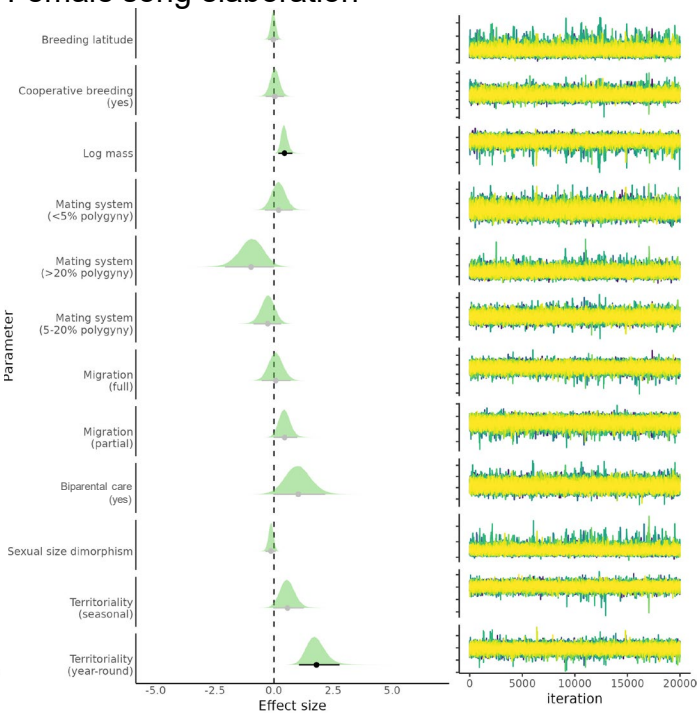


Fig. S2. Associations of predictor variables with female song ordinal scores and duets. Effect size plots with posterior distributions and trace plots for brms results of female song analyses. Black bars that do not cross zero on the effect size plot are variables that contribute substantially to each model.

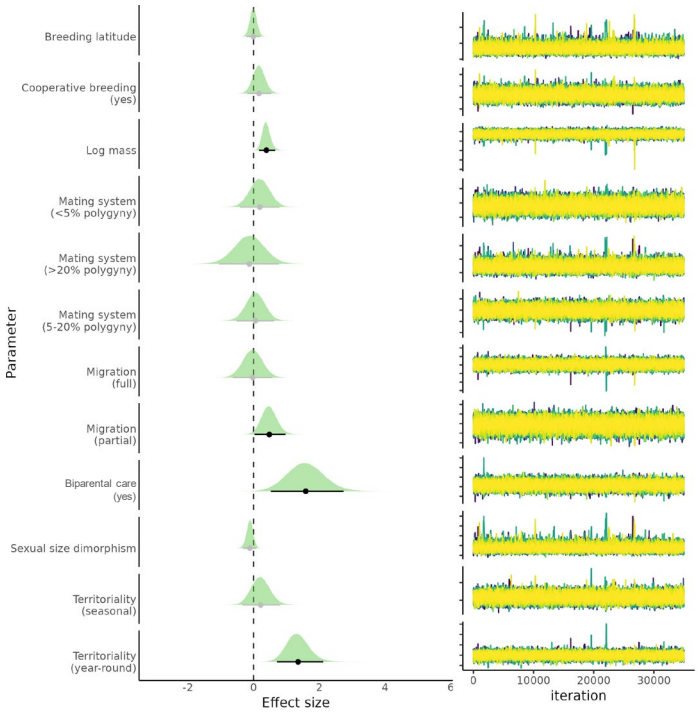
Female song incidence



Female song elaboration



Female song length



No female song vs Female solo song vs Duetting

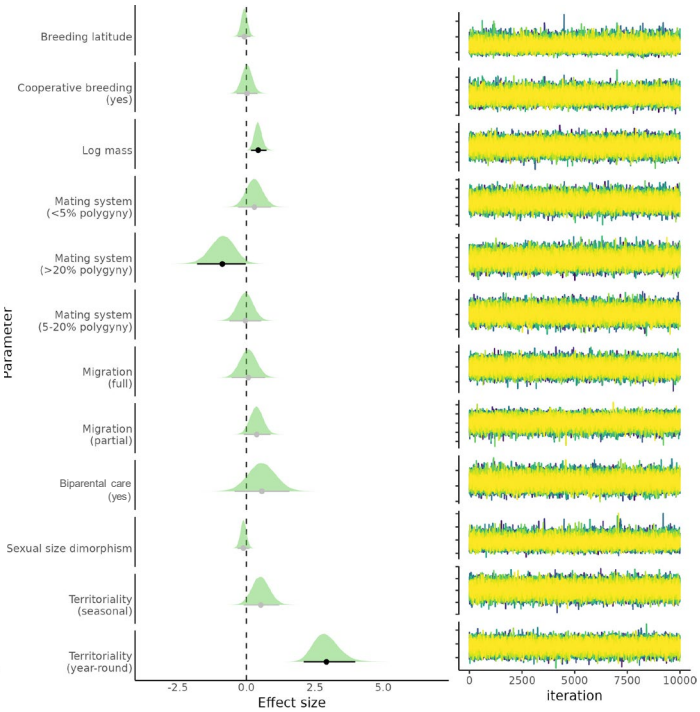


Fig. S3. Associations of predictor variables with female song present vs absent. Effect size plots with posterior distributions and trace plots for brms results comparing predictor variables to the binary response variable of female song present vs absent. Black bars that do not cross zero on the effect size plot are variables that contribute substantially to each model.

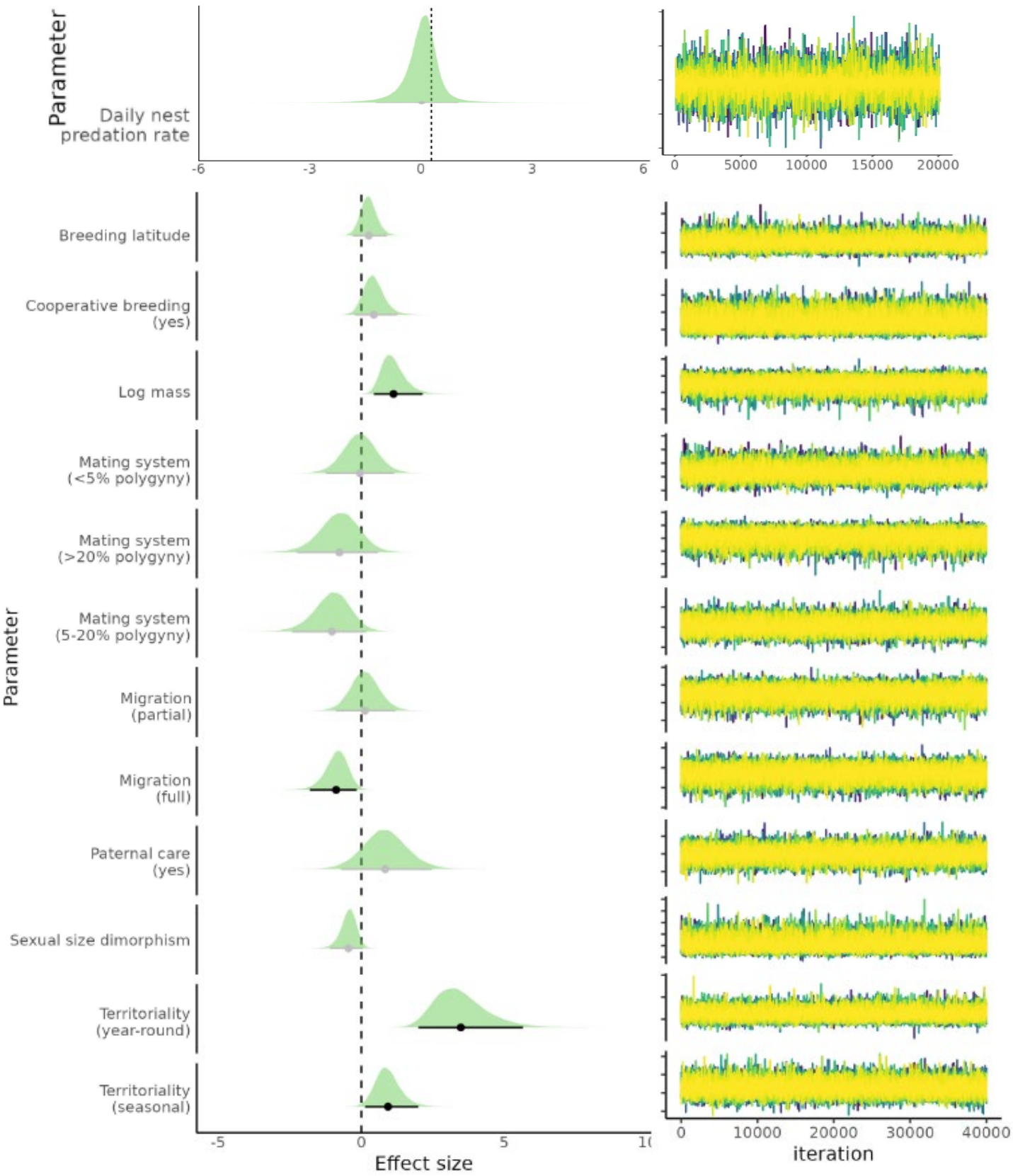
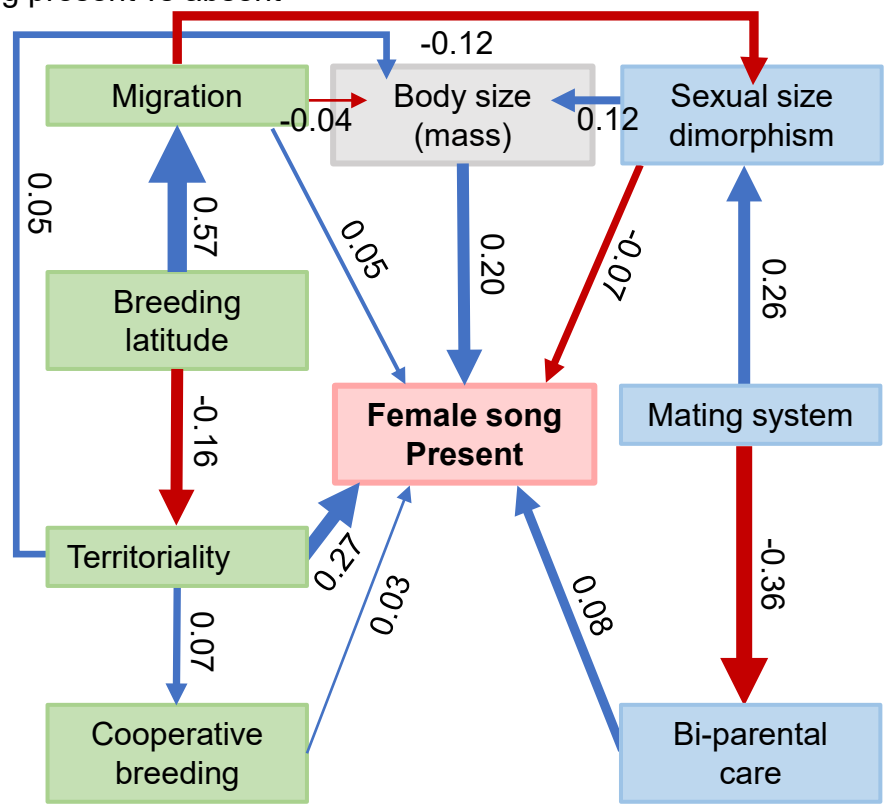


Fig. S4. Best models for phylogenetic path analysis for (a) all female song data coded as female song present vs absent and (b) ordinal incidence data with species without female song removed. Arrows represent supported paths among predictor variables. The thickness of each arrow represents the strength of each association, determined by the correlation coefficient next to each arrow. Blue arrows = positive correlations and red arrows = negative correlations. Box colours match hypotheses from Fig. 1.

(a) Female song present vs absent



(b) Female song incidence (species without female song Removed)

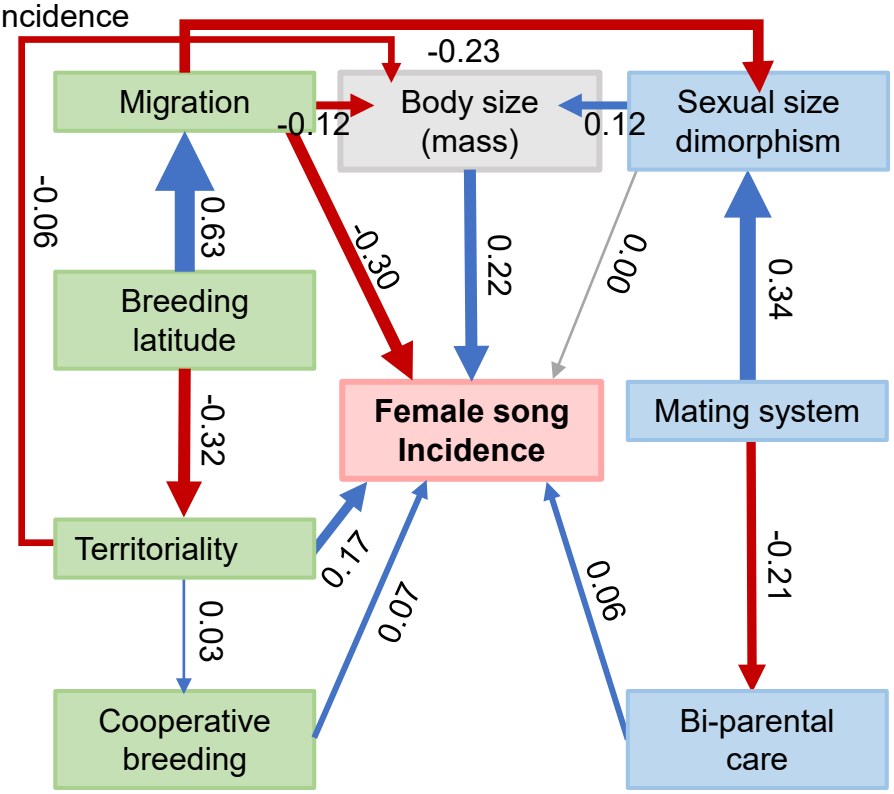
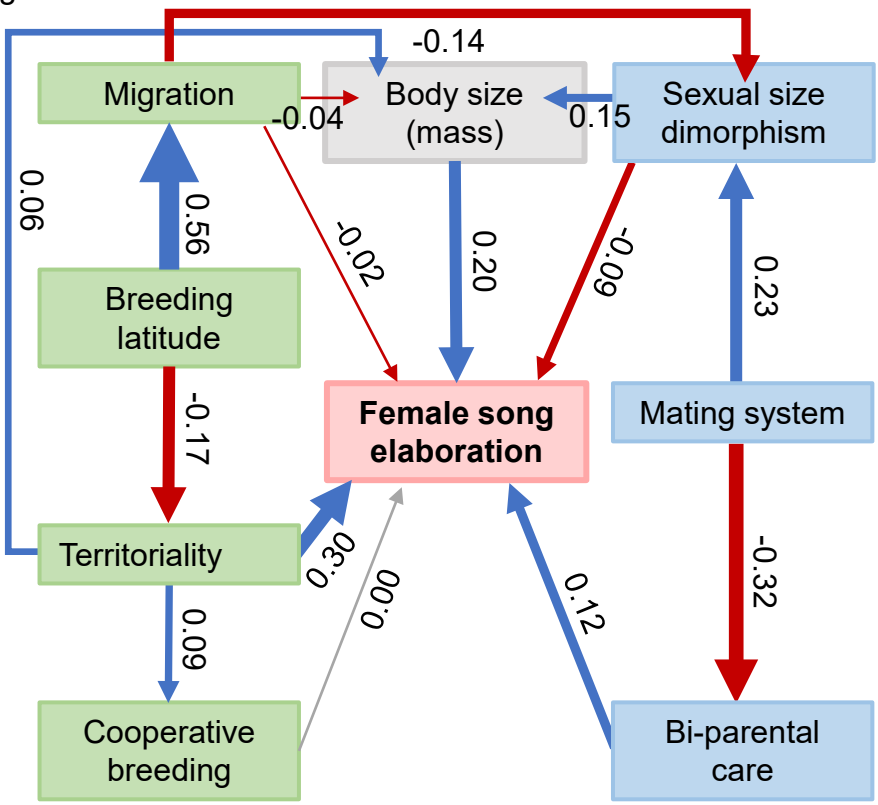


Fig. S5. Best models for phylogenetic path analysis for (a) song elaboration and (b) song length. Arrows represent supported paths (associations) among predictor variables. The thickness of each arrow represents the strength of each association, determined by the correlation coefficient next to each arrow. Blue arrows represent positive correlations and red arrows are negative correlations. Box colours match hypotheses from Fig. 1.

(a) Female song elaboration



(b) Female song length

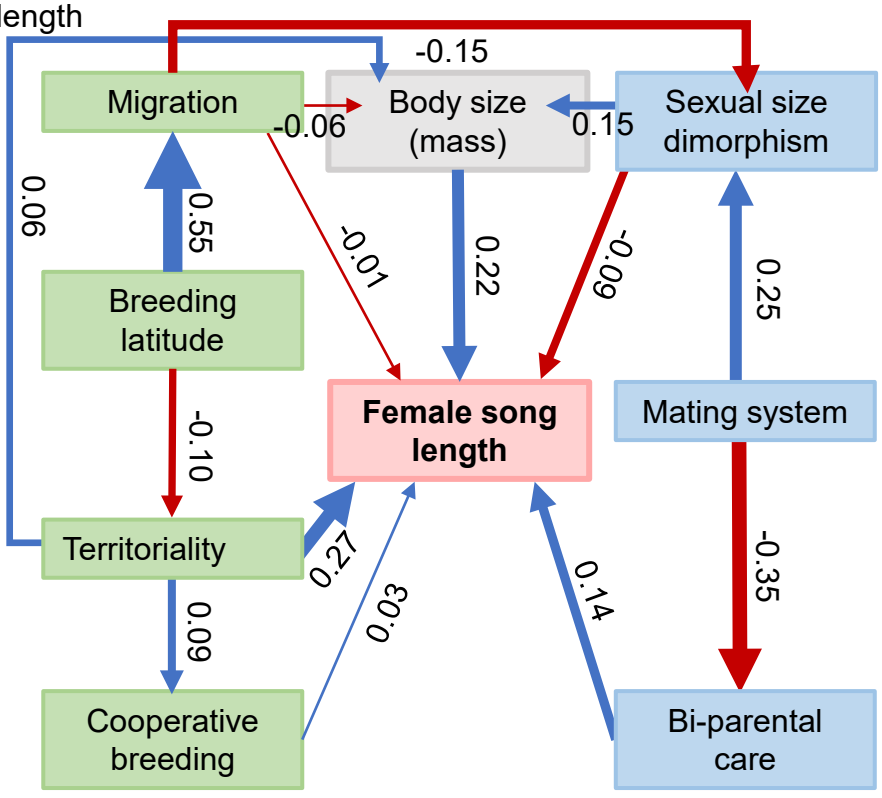


Fig. S6. Interaction of mating system with territoriality as it pertains to female song incidence. Effect size plots with posterior distributions and trace plots for brms results of female song analyses. Black bars that do not cross zero on the effect size plot are variables that contribute substantially to each model.

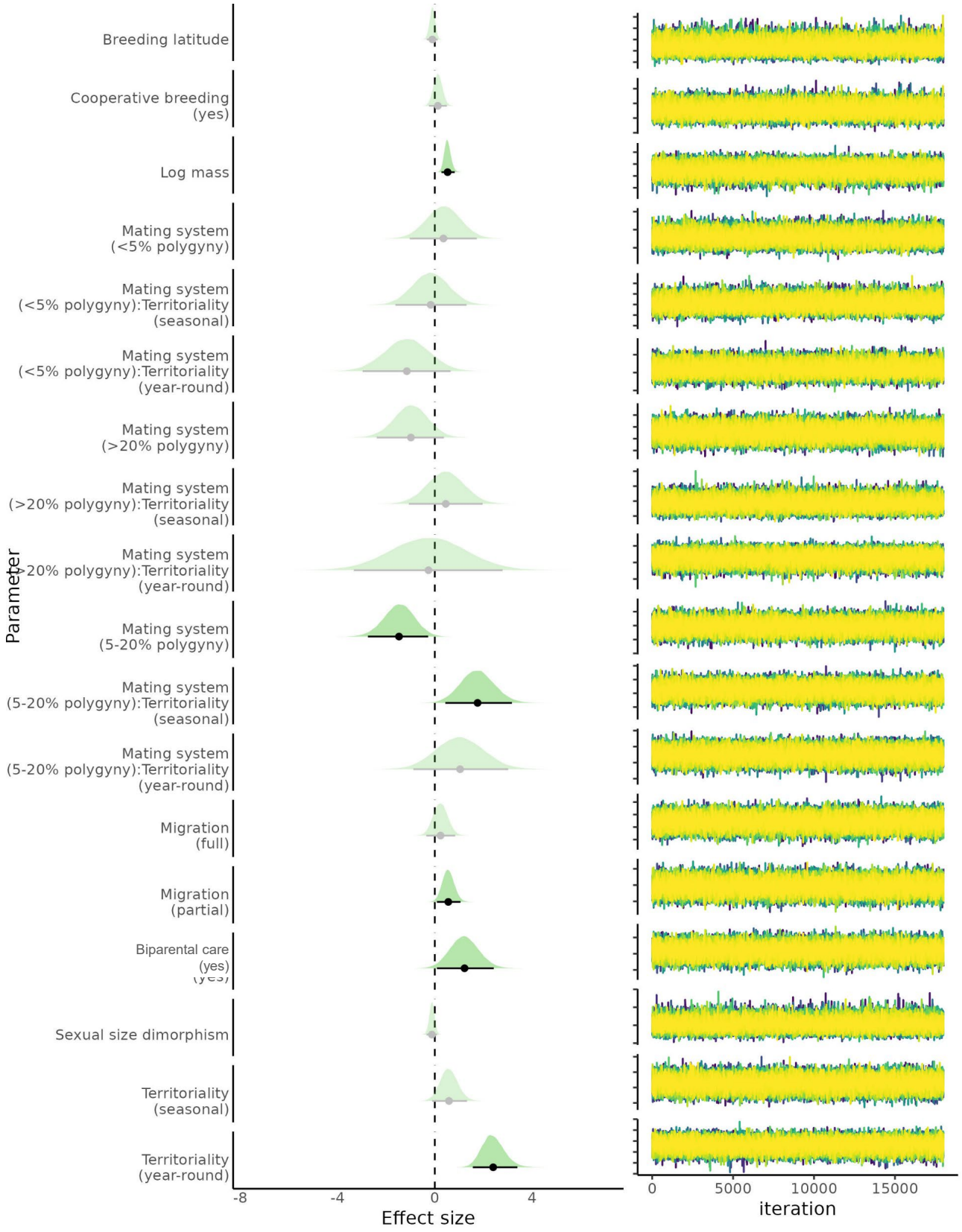
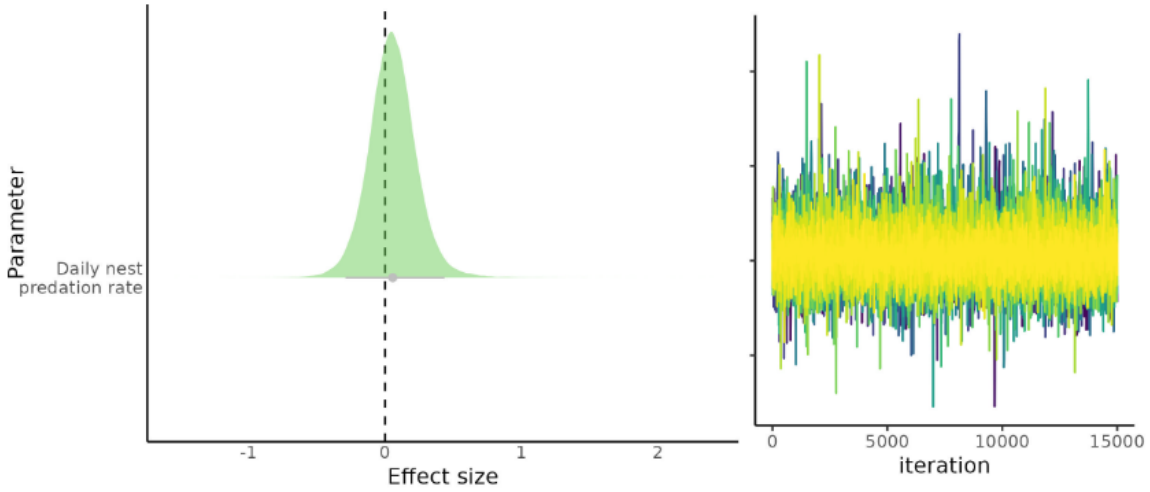
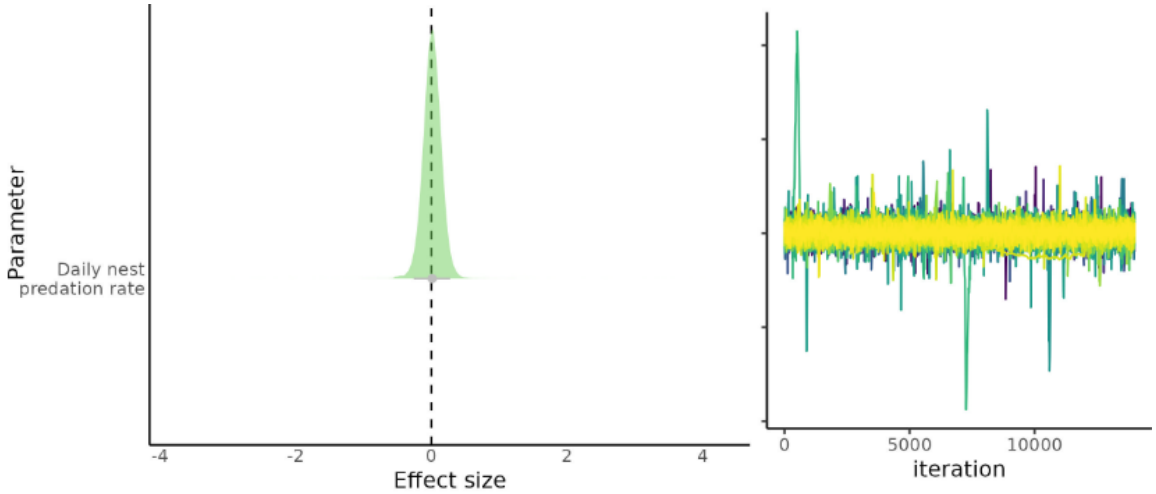


Fig. S7. Daily Nest Predation. Effect size plots with posterior distributions and trace plots for brms results of female song analyses. Black bars that do not cross zero on the effect size plot are variables that contribute substantially to each model.

Female song incidence



Female song elaboration



Female song length

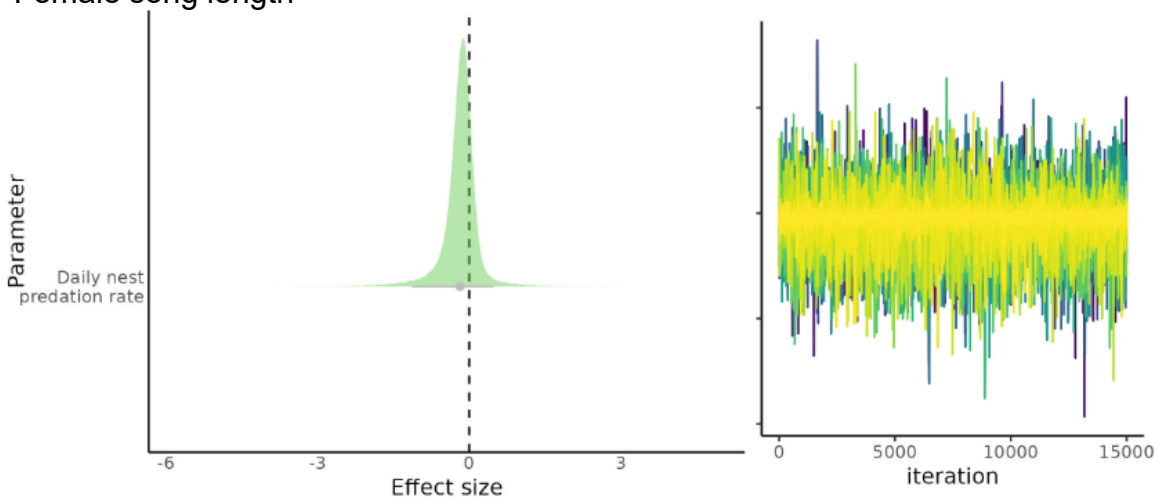
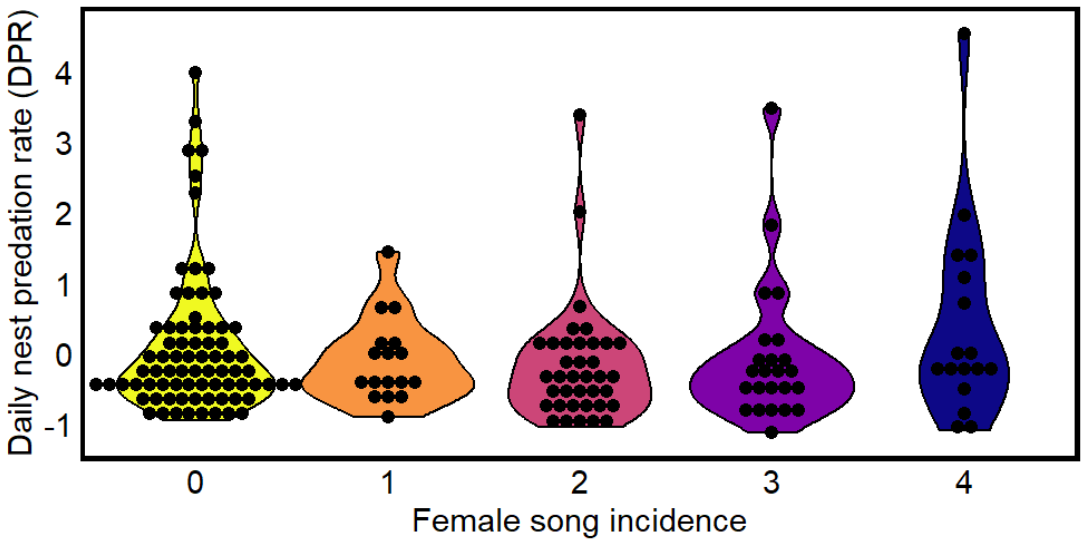
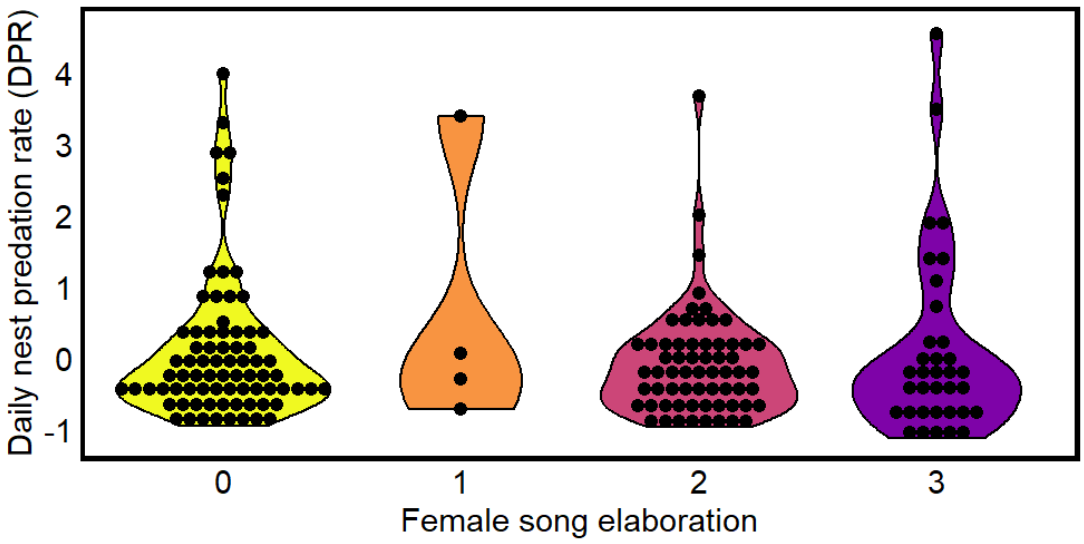


Fig. S8. Violin plots of the three female song metrics and daily nest predation rates. Black dots represent sample size for each category and the outlines represent the distribution of the data.





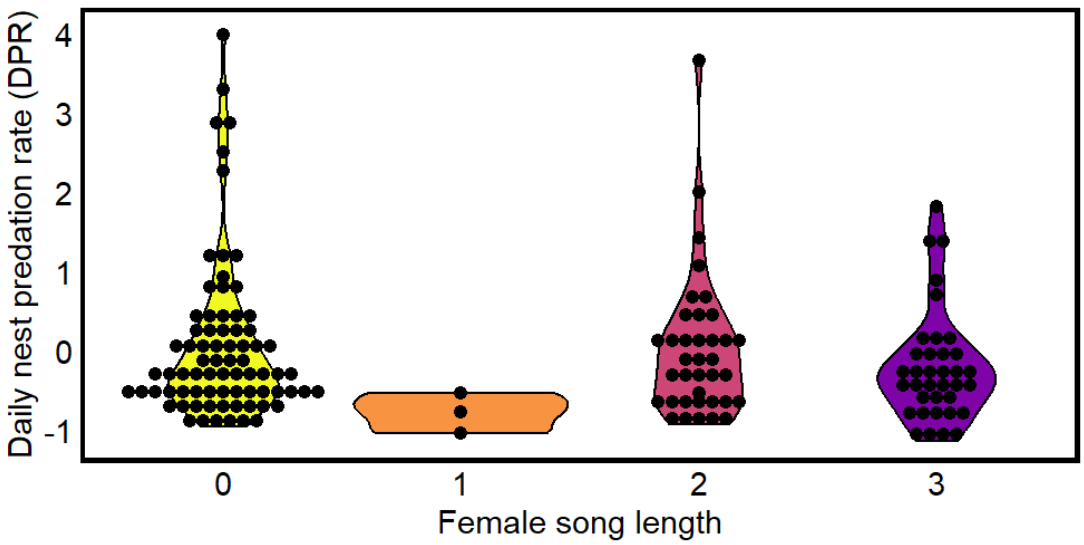


Fig. S9. (a) Female song incidence compared to duet (presence/absence) showing that species with high female song rates are also often duetting species. **(b-f) Female song incidence analyzed as an ordinal score from** (1) species without female song (absent), to (2) species with female solo song (non-duetting), and (3) species with female song that are also duetting species, compared to natural history traits.

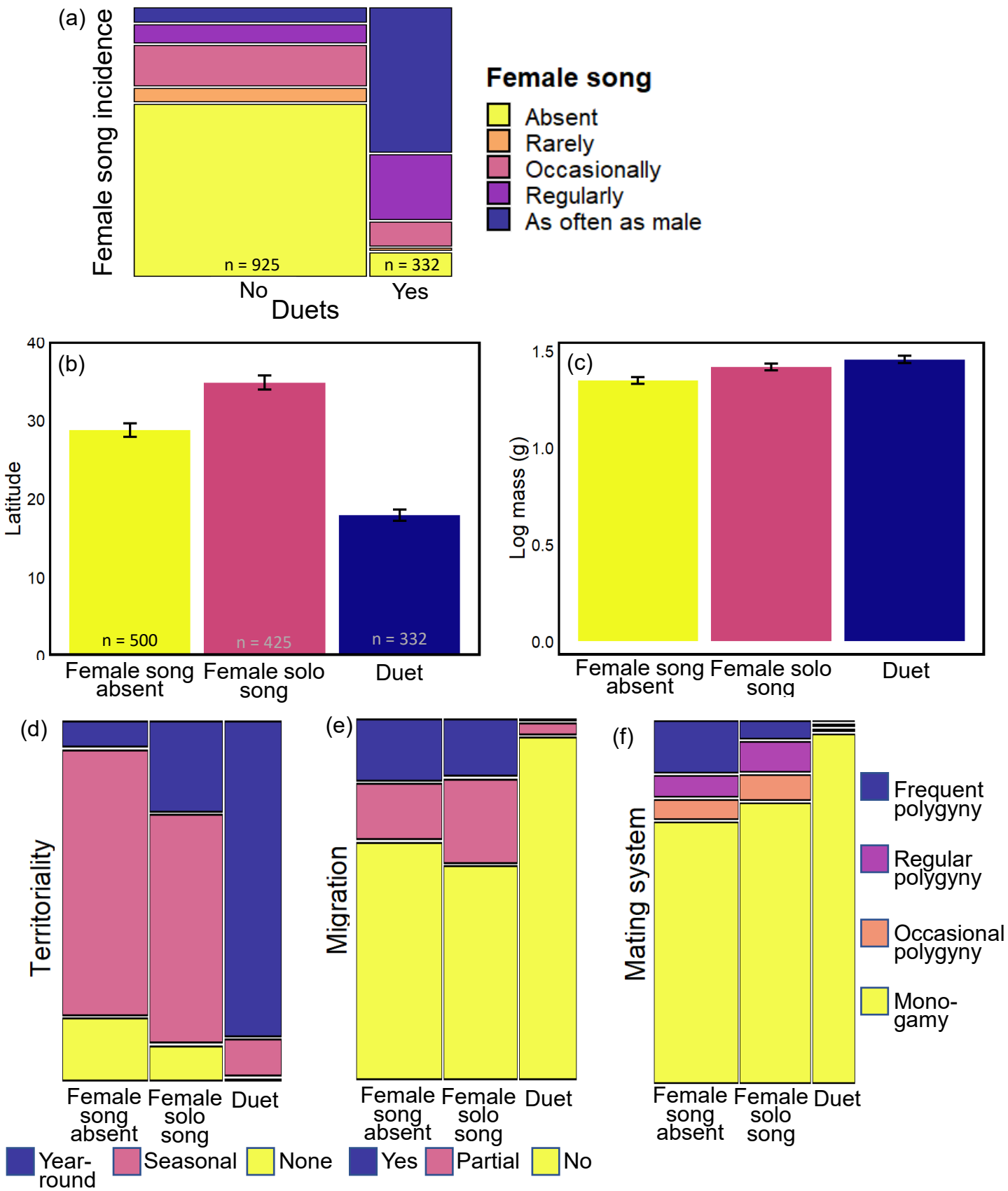


Fig. S10. Natural history attributes of species with female song (female song present) but unknown values for incidence (not enough information to score incidence). These data indicate that these species tend to breed at lower latitudes and be monogamous, year-round or seasonally territorial and non-migratory with biparental care. Some of these species also have low, moderate, or high rates of polygyny, no territoriality, migratory behaviour, and/or lack paternal care. Numbers overlaid on the graph are sample size (n) and the number of species with unknown incidence (u) for each category.

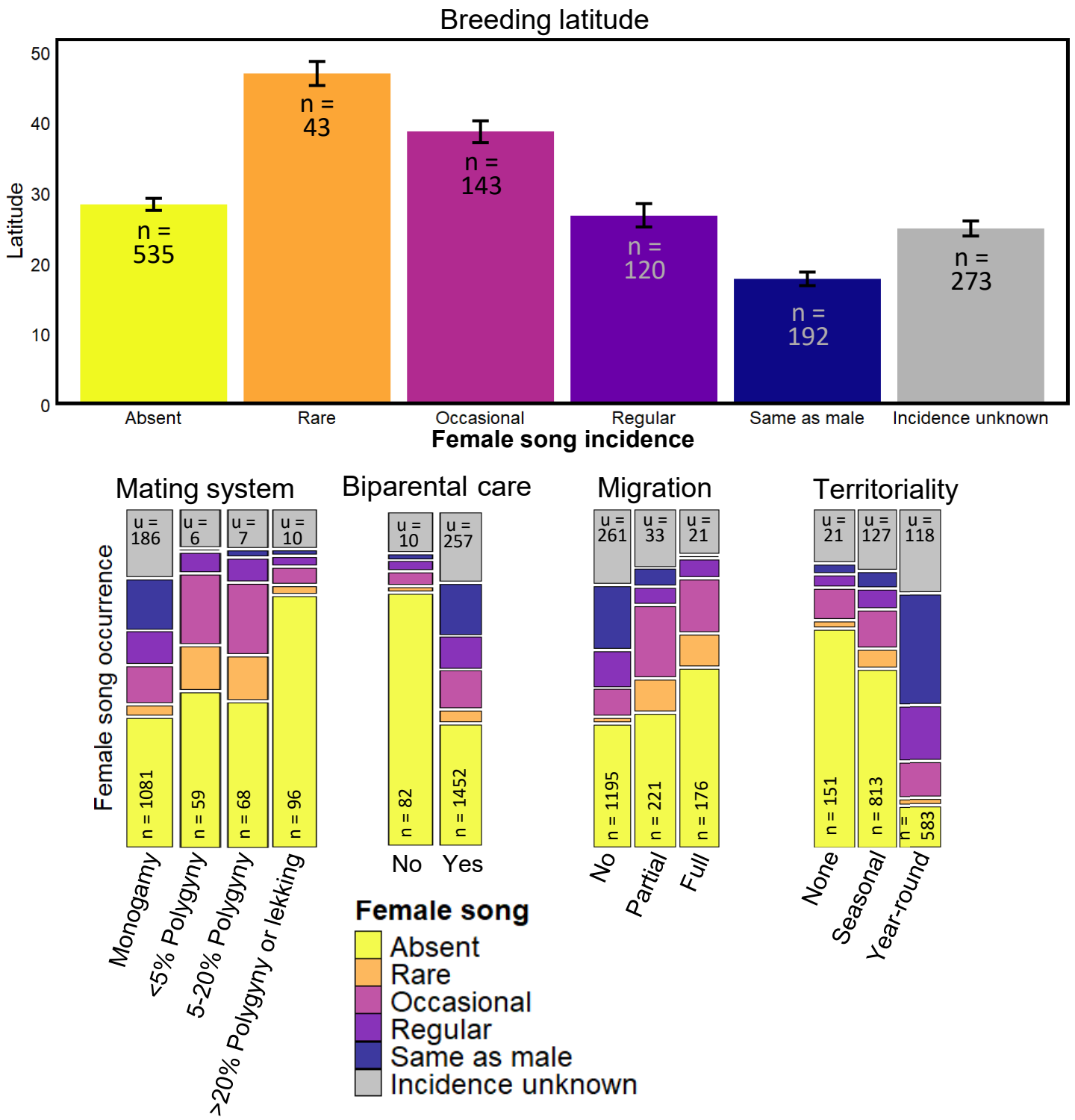
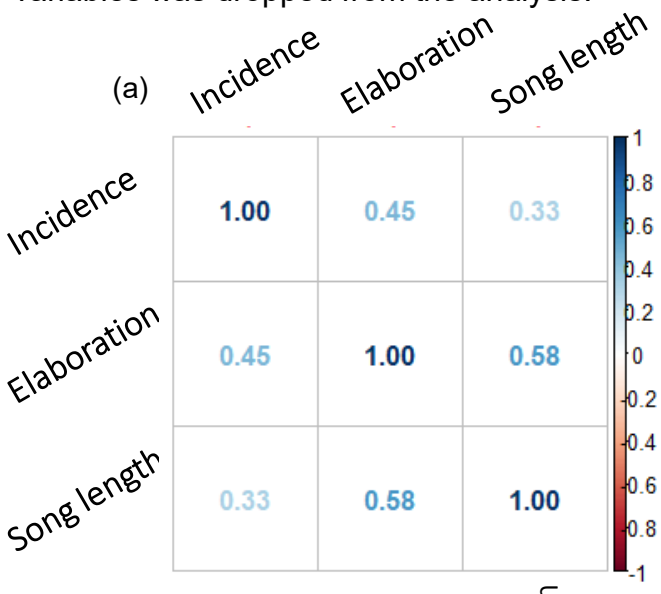


Fig. S11. Pearson correlations among (a) response variables and (b) potential predictor variables. For variables with a correlation coefficient higher than 0.7, one of the two variables was dropped from the analysis.



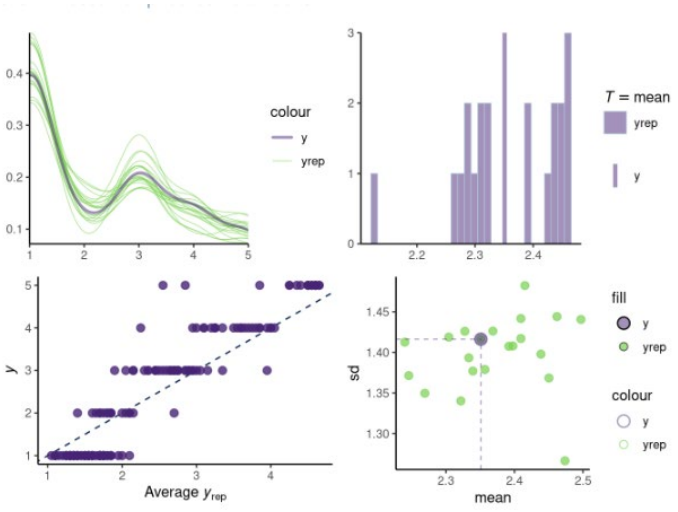
(b)

	Daily nest pred	Log mass	Log wing length	Size dimorphism	Mating system	Biparental care	Cooperative breeding	Migration	Log clutch size	Degrees from equator	Seasonality	Territoriality	Duet (Y/N)	Social bond length
Daily nest predation	1.00	-0.15	-0.17	0.00	-0.15	0.12	-0.20	-0.13	-0.06	-0.18	-0.13	-0.05	0.07	0.16
Log mass	-0.15	1.00	0.93	0.22	0.14	-0.32	-0.10	-0.20	-0.29	-0.09	-0.16	0.04	0.10	0.02
Log wing length	-0.17	0.93	1.00	0.20	0.11	-0.28	-0.11	-0.11	-0.22	-0.03	-0.09	-0.01	0.02	-0.03
Size dimorphism (wing)	0.00	0.22	0.20	1.00	0.42	-0.40	-0.09	0.01	-0.16	0.10	-0.01	-0.14	0.04	-0.12
Mating system	-0.15	0.14	0.11	0.42	1.00	-0.43	-0.03	0.29	0.12	0.22	0.18	-0.32	-0.20	-0.52
Biparental care	0.12	-0.32	-0.28	-0.40	-0.43	1.00	0.12	0.07	0.28	0.06	0.13	0.15	0.02	0.24
Cooperative breeding	-0.20	-0.10	-0.11	-0.09	-0.03	0.12	1.00	-0.12	0.02	-0.03	0.08	0.24	0.12	0.19
Migration	-0.13	-0.20	-0.11	0.01	0.29	0.07	-0.12	1.00	0.49	0.62	0.60	-0.41	-0.35	-0.63
Log clutch size	-0.06	-0.29	-0.22	-0.16	0.12	0.28	0.02	0.49	1.00	0.57	0.61	-0.30	-0.28	-0.30
Degrees from equator	-0.18	-0.09	-0.03	0.10	0.22	0.06	-0.03	0.62	0.57	1.00	0.85	-0.25	-0.18	-0.46
Seasonality	-0.13	-0.16	-0.09	-0.01	0.18	0.13	0.08	0.60	0.61	0.85	1.00	-0.26	-0.25	-0.42
Territoriality	-0.05	0.04	-0.01	-0.14	-0.32	0.15	0.24	-0.41	-0.30	-0.25	-0.26	1.00	0.46	0.48
Duet (Y/N)	0.07	0.10	0.02	0.04	-0.20	0.02	0.12	-0.35	-0.28	-0.18	-0.25	0.46	1.00	0.38
Social bond length	0.16	0.02	-0.03	-0.12	-0.52	0.24	0.19	-0.63	-0.30	-0.46	-0.42	0.48	0.38	1.00

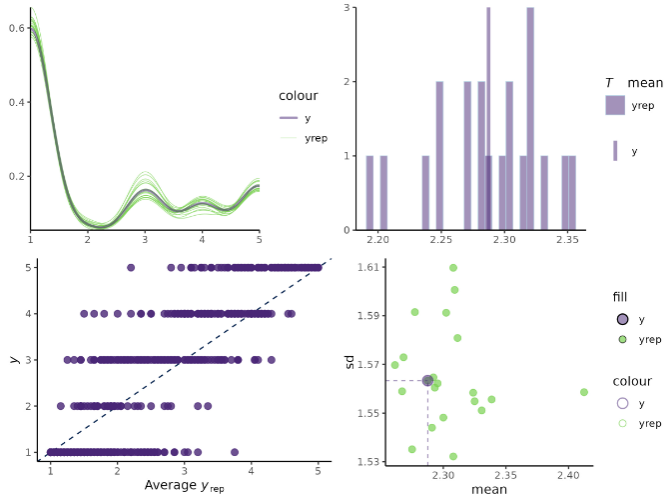
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1

Fig. S12. Posterior predictive check plots to assess model fit for female song incidence and female song vs duets. Similar plots were created for all analyses and all showed similarly good fits. An explanation of each set of four plots (per analysis), is as follows: The top left panel shows the density plot comparing the observed data (y, purple line) with multiple posterior predictive simulations (yrep, green lines). This plot visualizes how well the model captures the shape of the observed data distribution. The top right panel presents a histogram of the mean values; the dark purple bar represents the mean of the observed data, and the light purple bars display the distribution of means from the posterior predictive simulations, illustrating variability in the model's predictions. The bottom left panel is a scatter plot showing the relationship between the observed data (y, purple dots) and the average of the simulated data (yrep). The dashed line indicates where the observed data aligns with the model average prediction, with points closer to this line showing better model agreement. The bottom right panel shows the mean against the standard deviation for the observed data and the predicted values of the model. The large purple dot represents the mean and standard deviation of the observed data, while the green dots correspond to the same statistics from the predicted data. This panel helps assess whether the model captures the central tendency and variability of the data.

Female song incidence: nest predation model



Female song incidence: natural history model



Female song absent/solo/duet vs natural history model

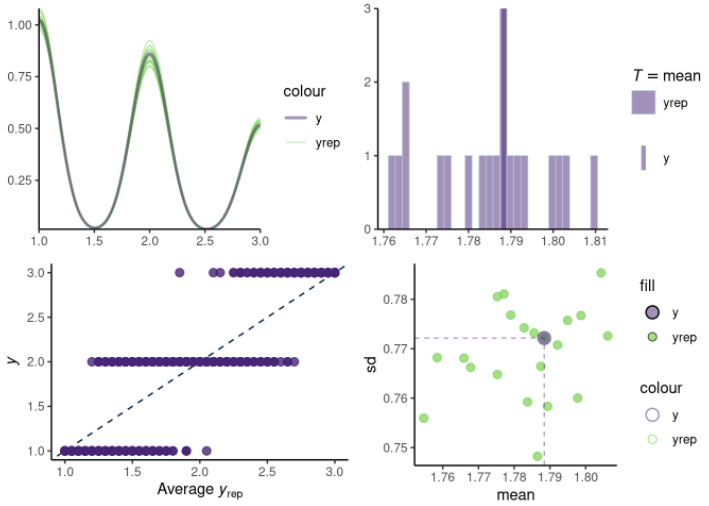


Fig. S13. Subset of hypotheses tested via phylogenetic path analysis. Below are the full tested interactions among the predictor variables for each major hypothesis. We examined these main hypotheses independently, as well as combinations of interactions within and among the hypotheses. In addition, we sequentially dropped the direct connection between each predictor variable and the response variable to evaluate direct effects of our predictor variables on female song. We tested the same full set of models for each female song metric.