Territoriality modulates the coevolution of cooperative breeding and female song in songbirds

Kate T. Snyder[1,2], Aleyna Loughran-Pierce[1,3,4], Nicole Creanza[1,2,*]

[1] Department of Biological Sciences, Vanderbilt University, Nashville, TN, 37212
[2] Evolutionary Studies Initiative, Vanderbilt University, Nashville, TN, 37212
[3] School for Science and Math at Vanderbilt, Vanderbilt University, Nashville, TN, 37212
[4] Current address: Department of Environmental Sciences, University of Virginia, Charlottesville, VA, 22904

[*] Corresponding author
Email: nicole.creanza@vanderbilt.edu (NC)

Preprint: https://doi.org/10.1101/2024.06.26.600822

Data and code (private repository for peer review): https://figshare.com/s/7e9da2429ee8e50c279d

Keywords: birdsong, learned behaviors, cultural evolution, female song, social behavior

Abstract

Birdsong has historically been characterized as a sexually selected, primarily male behavior. More recent findings suggest female song is widespread, raising questions about how social functions of birdsong shape song evolution. Certain breeding systems, like cooperative breeding, change social dynamics and selection pressures on both sexes, providing important contexts for studying song evolution. Here, we use phylogenetic comparative analyses across 1041 songbird species to examine relationships between cooperative breeding, female song, and male song characteristics. Here, we show robust bidirectional co-evolutionary dynamics between cooperative breeding and female song that persist when controlling for territoriality, allometry, and geographic sampling biases. Importantly, when examining intensity of territorial defense, we find this relationship is context-dependent: while cooperative breeding and female song commonly co-occur in strongly territorial systems, their association is especially pronounced in weakly territorial systems, where they co-occur much more often than expected by chance. Additionally, we observe that male song repertoire size evolves more slowly in cooperative-breeding lineages. These findings demonstrate that cooperative breeding shapes the evolution of vocal communication differently based on territorial context and sex, with female song potentially serving crucial but understudied functions related to social cohesion in cooperative systems, particularly in species where territorial conflict is reduced.

Introduction

Song is traditionally viewed as a sexually selected male behavior with functions related to mate attraction and territory defense[1], but increased recognition that female songbirds sing has revealed that this framing may be incomplete[2,3]. Data on the biological contexts of female song are sparse for most species, but, broadly, several patterns of singing behavior have been observed, including male-female duetting, agonistic female-female interactions, and within-pair or within-group interactions[3]. These behavioral patterns correspond with various hypothesized functions of female song: mate attraction, intrasexual competition, territory defense, pair-bond maintenance, and social cohesion within groups[4,5]. A recent quantitative review reinforces that territory defense, intrasexual competition, and intrapair communication are the most commonly supported functions of female song, while its use in mate attraction appears rare[6]. The evolution of song may thus be subject to evolutionary pressures related to social interaction in addition to traditionally studied sexual selection, a phenomenon that might be especially relevant in females[7,8].

These varied functions of female song appear to be shaped by ecological and life history factors. In particular, certain breeding systems, like cooperative breeding, change social dynamics and selection pressures on both males and females, providing important contexts for studying interactions between perturbed selection pressures and song evolution. Emerging evidence continues to highlight the social complexity of cooperative breeding systems, with recent work documenting cryptic reciprocal helping relationships between kin and non-kin that persist across breeding seasons[9]. Such durable social affiliations likely require sophisticated communication systems for coordinating helping behaviors, recognizing group members, and maintaining social bonds, and these communication systems might be more valuable to a group with shared resources to defend, or easier to maintain when the group has a defined territory. Understanding how cooperation and territoriality interact to shape vocal evolution is therefore important for explaining the diversity of singing behaviors across species.

Cooperative breeding, which is broadly defined as a breeding system in which non-parental adults assist in rearing young, has evolved numerous times in the avian lineage, and is observed to different extents in approximately 13% of Oscine songbird species (Passeri)[10]. In cooperative systems, sexual selection pressures on females are predicted to increase through intensified intrasexual competition for breeding status

and associated increase in variance of reproductive success among females[11,12], although the extent of this reproductive skew can vary widely by species, especially in avian taxa[13]. The effects of cooperative breeding on sexual selection pressures in males are even less straightforward[14,15]. Cooperative breeding often arises from kin-based social structures, with both inclusive and direct fitness benefits for group members, such as resistance to predation through increased group size[16–19]. Notably, cooperative breeding often co-occurs with territoriality and stable social bonds[20]. The fitness benefits of cooperative breeding, therefore, are somewhat entwined with other forms of group living. In songbirds, how these complex social systems affect the functions of and selection pressures on male and female song remains poorly understood. However, emerging evidence suggests that there may be important evolutionary interactions between social stability and vocalizations[21]; for example, cooperative breeding appears to promote the evolution of larger repertoires of functional vocalizations[22]. In addition, several traits related to social network stability, including territoriality, monogamy, and long-term pair bonds, are correlated with the presence of female song in various songbird clades[23]. Further, recent phylogenetic analyses[24] found that year-round territoriality and biparental care are the strongest predictors of female song incidence across songbirds, suggesting that ecological factors related to resource defense and parental investment shape vocal evolution. Notably, this study found an additional, albeit weaker, persistent link between cooperative breeding and female song, despite cooperative breeding being considered a subset of biparental care. Overall, there are hints that living in social groups, defending territories, and engaging in cooperative caretaking might influence the evolution of female and male vocalizations, but the intricacies of these links have not been fully disentangled.

Here, we compile a dataset of Oscine species that includes species-level cooperative breeding classification, data on female song presence, several features of (putatively male) song complexity and performance, and novel classifications of territorial intensity. Using phylogenetically informed analyses across 1041 species, we test whether cooperative breeding influences the evolution of male song traits or female song presence, and how territoriality shapes those interactions. We also evaluate other social variables to compare the effects of cooperative caretaking with other factors related to group living and social stability. Our analyses reveal that the evolutionary association between cooperative breeding and female song is both robust and context-dependent, with an outsized effect in weakly territorial systems. In contrast, male song

repertoire size evolves more slowly in cooperative breeding lineages regardless of territorial context. These sex-specific, context-dependent patterns suggest that multiple evolutionary pathways could promote female song: resource defense in territorial species and social coordination or cohesion in cooperative contexts.

Results

Here, we synthesize four novel datasets and analyze them together with large-scale avian phylogenies. First, we updated our previously published dataset on species-level song features encompassing song complexity and performance metrics for 339 species (S1 Dataset). Second, we combined and reconciled data sources reporting the species-level prevalence of cooperative breeding for 4373 Oscine species (S2 Dataset). Third, we combined and updated datasets on the presence or absence of female song, obtaining classifications for 1094 species (S3 Dataset). Fourth, we refined existing territorial behavior classifications[21] by developing novel designations of weak or absent territoriality versus strong territoriality for 875 species based on detailed behavioral criteria (S4 Dataset). We also compiled species-level data[16,25] on other forms of social organization (Tables S1-2). Together, these datasets allow us to elucidate the complex relationships between territorial behavior, cooperative breeding, and the evolution of song in both sexes.

Song feature evolution in the context of cooperative breeding and other sociality characteristics

We found that none of the species-level song features were significantly different between non-cooperative and cooperative breeding species (phylANOVA $p > 0.05$, Table S3). We did, however, find that the rates of evolution of two song metrics associated with song complexity, syllable repertoire size and song repertoire size, may have differed between non-cooperative and cooperative lineages, with both evolving more rapidly in non-cooperative lineages in over 95% of simulations, though this difference was only significant for song repertoire size (Brownie likelihood-ratio test: song repertoire $p = 0.024$, n = 209 species; syllable repertoire $p = 0.136$, n = 118 species; Fig. 1, Table S4). We obtained qualitatively similar results for song repertoire evolution when using variations of the cooperative breeding or song repertoire data across cited

sources, when using variations of the phylogeny to account for phylogenetic uncertainty, or when performing a jackknife analysis by removing each family in turn (Tables S4,S5, Extended Data Fig. 1 [ED1]).

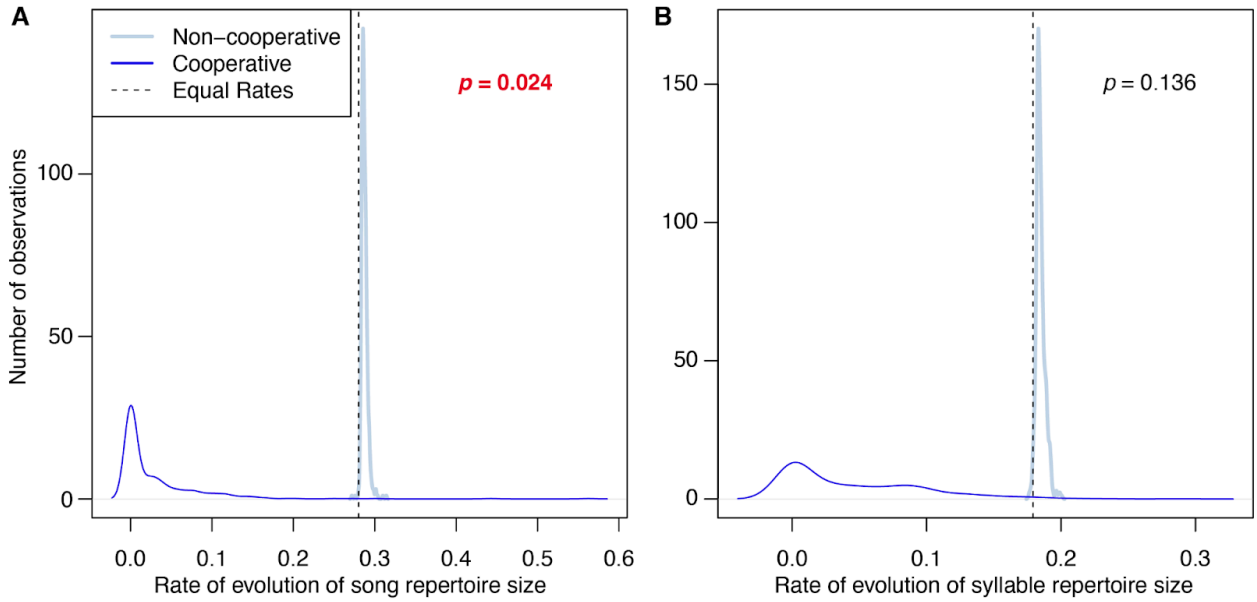


Fig. 1: Testing whether the rate of evolution of (A) song repertoire and (B) syllable repertoire (both natural-log-transformed) differs based on cooperative breeding. Song repertoire evolved faster in non-cooperative breeding lineages in 99.2% of simulations, with a Brownie likelihood-ratio test $p = 0.024$ (n = 209 species). Syllable repertoire showed a similar trend of faster evolution in non-cooperative breeding lineages in 98.8% of simulations of the Brownie algorithm[26] (Brownie likelihood-ratio test $p = 0.136$, n = 118 species). Both analyses were run for 500 simulations.

We similarly found no significant song repertoire size differences between species in different states of any of the discrete sociality traits (phylANOVA $p > 0.05$, Table S3). However, in our analyses of evolutionary rates, we found that song repertoire likely evolved faster in lineages with non-familial versus familial living (Brownie likelihood-ratio test $p = 0.0042$, n = 156 species) and faster in lineages with social bonds lasting one breeding season or less ($p = 0.0079$, n = 129 species) (Table S4). We also found that song repertoire size evolved fastest in lineages with weak, seasonal, or absent territoriality (territory: binary $p = 0.011$, three-level $p = 0.012$, n = 216 species; Fig. ED2B, Tables S4,S6B). The analysis comparing song repertoire evolution between lineages with each combination of cooperative breeding and binary territoriality states (non-cooperative breeding with weak territoriality, non-cooperative with strong territoriality, cooperative breeding with weak territoriality, and cooperative breeding with strong territoriality) showed that song repertoire evolved more slowly in both cooperative states than in either non-cooperative state, but that non-cooperative weakly territorial lineages had faster song repertoire evolution than non-cooperative strongly territorial lineages

(p = 0.011, n = 152 species; Fig. ED2C, Table S6C). In [16], social-system data for 1203 species were classified into either three or four states; our analyses showed that song repertoire evolved more rapidly in non-familial, non-cooperative lineages than in lineages with familial living plus cooperative breeding, familial living without cooperative breeding, or non-familial living plus cooperative breeding (Brownie likelihood-ratio test: three-state p = 0.002, four-state p = 0.001, n = 156 species; Fig. ED2D,E, Table S6D-E), recapitulating our findings from analyzing the two traits separately, namely that song repertoire size evolves more slowly both in familial lineages and in cooperative lineages (Table S4).

Female song and cooperative breeding

We were able to obtain species-level classifications of both female song presence/absence and cooperative breeding for 1041 Oscine species, and compared their evolution using the assembled species-level data as character states at the tips of a published avian phylogeny[27] (Fig. 2). By simulating evolution of these traits[28], we found evidence to reject the null hypothesis that female song and cooperative breeding evolve independently (empirical p = 0.002), a finding primarily driven by a significantly higher-than-expected co-occurrence of female song presence and cooperative breeding (Fig. 3A, Table ED1). This trend in ancestral state association persisted when we iteratively removed taxonomic families (Table S7), and when we used a set of randomly sampled phylogenies, a consensus phylogeny calculated using an alternate method, and alternative cooperative breeding classification methods (Fig. ED3, Tables S8, S9).

By independently simulating transitions between breeding systems and female song states, we can tease out some potential directionality in the evolutionary patterns underlying this relationship. We found that transitions from presence to absence of female song were less frequent in cooperative breeding lineages than expected, suggesting that cooperative breeding might slow the loss of female song (Fig. 3B). In contrast, transitions from presence to absence of female song were more frequent than expected in non-cooperative breeding lineages. These trends were consistent when we ran this analysis using alternate data or phylogenies (Fig. ED4). Additionally, the presence of female song may promote a transition to cooperative breeding (Fig. 3B), though this trend was not as robust to alternative cooperative breeding classification methods (Fig. ED4E-I).

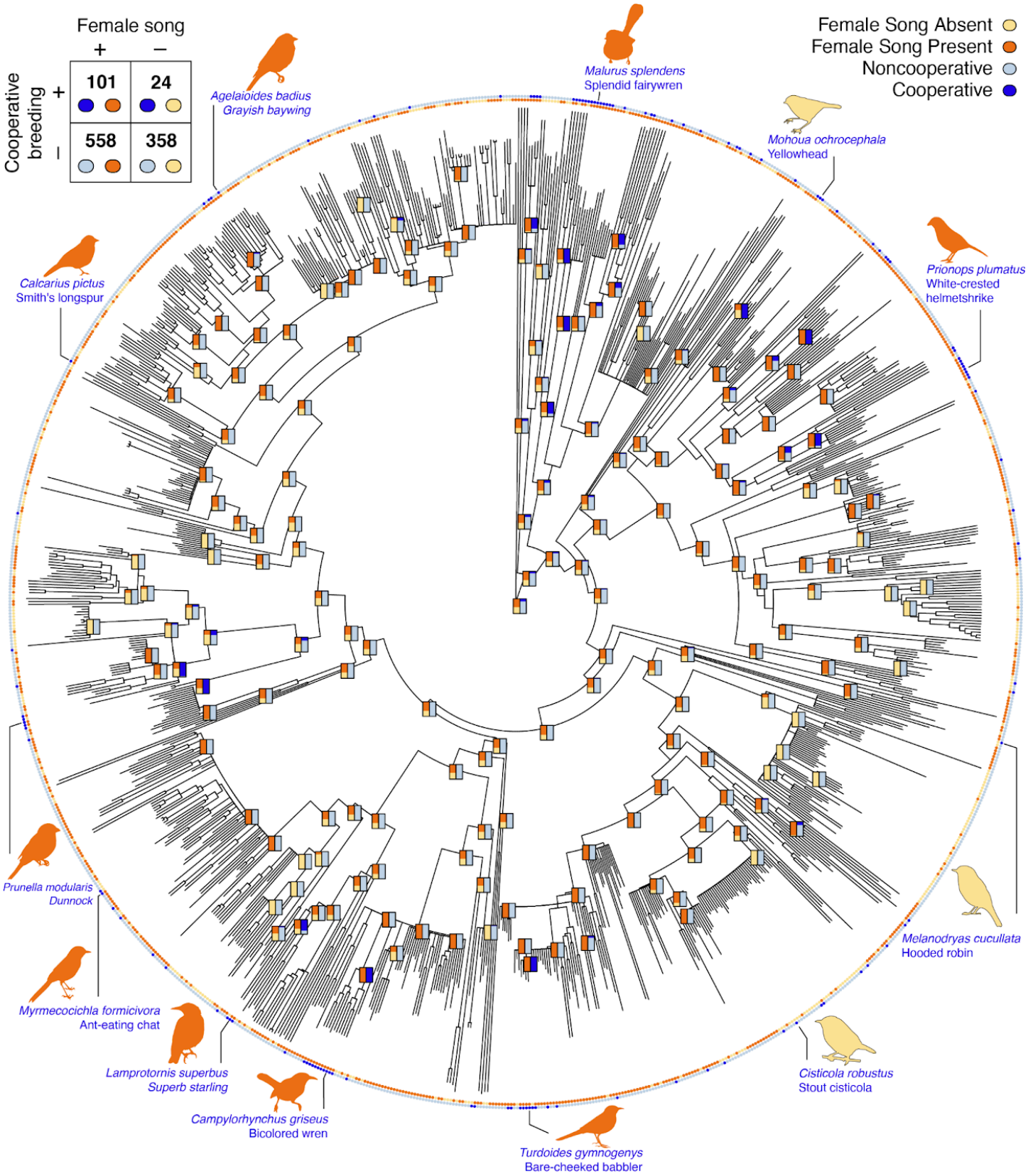


Fig. 2: Phylogeny showing the classifications of cooperative breeding (outer circle) and female song (inner circle) for 1041 Oscine species. The table in the top left corner shows the number of species in each category, with female song

overrepresented in cooperative breeding species. As landmarks, we illustrate several well-studied species that breed cooperatively, with their silhouettes colored according to whether female song is present in the species. We note that there are multiple evolutionary gains and losses of both cooperative breeding and female song. In general, the ancestral nodes that are estimated to lack cooperative breeding (light blue) also lack female song (yellow), and the ancestral nodes that are estimated to have cooperative breeding (dark blue) also usually have female song (orange).

Accounting for other social variables

Several social traits co-occurred with cooperative breeding more often than expected in evolutionary history: social living in groups smaller than about 30 individuals, social bonds longer than one breeding season, familial living, acolonial living, social monogamy, year-round and strong territoriality, and more than two caretakers (Table S10). Despite the high degree of co-occurrence between familial living and cooperative breeding, familial living and female song did not co-occur significantly more than expected ($p = 0.092$, n = 468 species, Fig. ED5A, Table ED1). A similar analysis using the four-state breeding system classification dataset showed that female song occurs significantly more often in systems with kin-cooperative breeding than expected and significantly less often in systems with non-kin-cooperative or non-familial, non-cooperative systems (Fig. ED5B, Table S11).

The only social variable that significantly co-occurred with both cooperative breeding and female song was territoriality—defined as either territory strength or year-round territoriality (Table S10, Table ED1). However, an initial assessment of species counts suggests that this mutual association with territoriality only partially explains the link between female song and cooperative breeding (Table S12). We thus included these territoriality variables in all subsequent analyses to thoroughly test whether the association between female song and cooperative breeding is explained by their mutual association with territoriality. Simulated transition count analyses split by territoriality state suggested that cooperative breeding might slow the loss of female song primarily in strong territoriality, and non-cooperative breeding might accelerate the loss of female song primarily in weak territoriality (Fig. 3C,D).

To further probe the interactions between female song, cooperative breeding, and territoriality, we also performed phylogenetic path analyses on networks that included these three traits, as well as species body mass (log-transformed) to account for allometric scaling effects (Fig. S1). We found that the models with the highest support ($\Delta CICc < 2$) always included a link between cooperative breeding and female song, though the

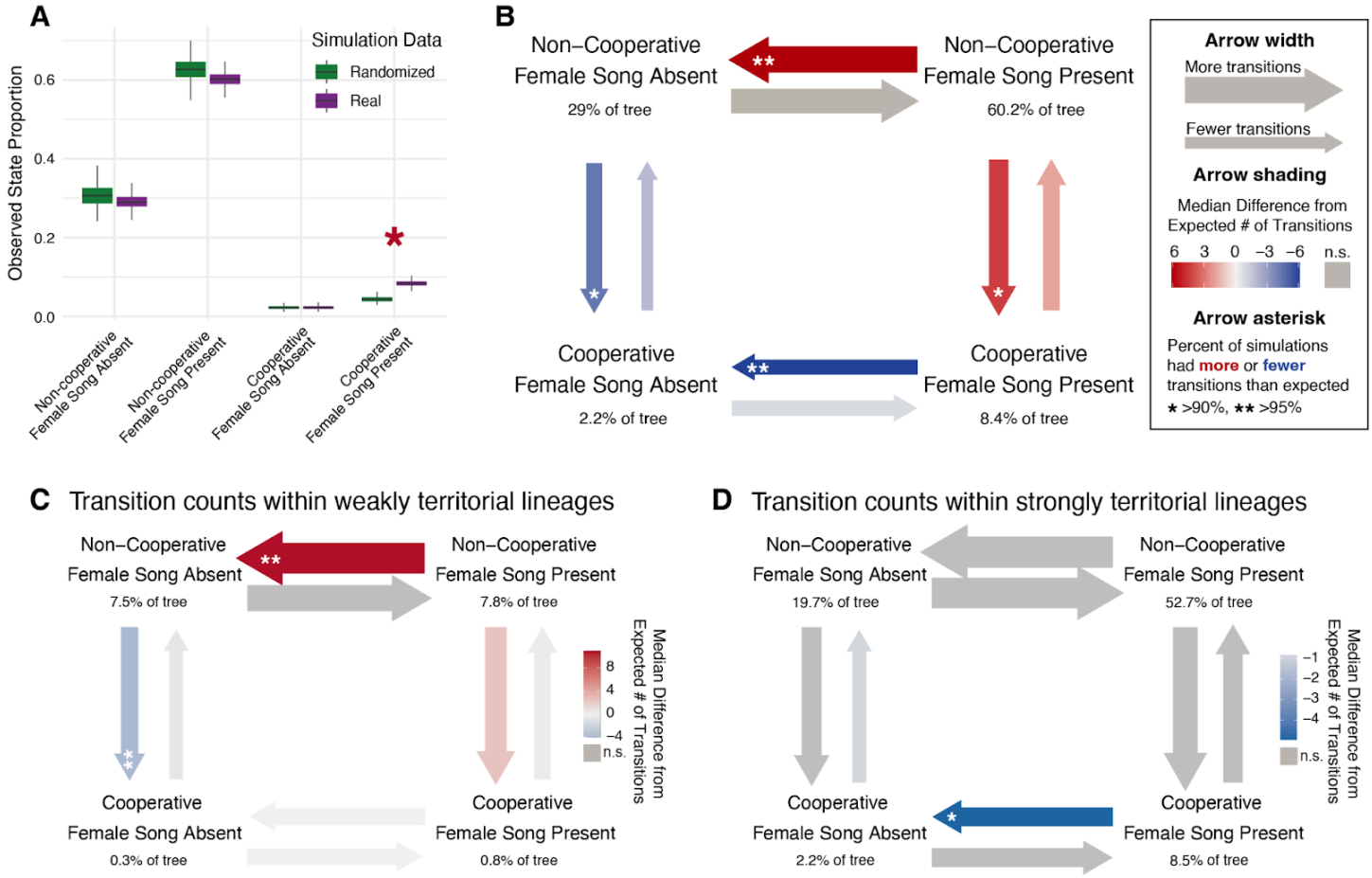


Fig. 3: Coevolution and correlated evolution of female song presence and cooperative breeding. A) Overlap of stochastic character map simulations of female song and cooperative breeding, measured as the proportion of the tree occupied by each state combination (empirical *p*-value 0.002, 500 simulations, N = 1041 species). B) Transition likelihoods between female song and cooperative breeding based on the median difference from expected transition counts from 500 simulations. Arrow weights scale to the log-transformed median relative number of state switches that occurred on that topology. Gray arrows indicate rates for which the observed number of transitions were not significantly different from the expected number, based on an ANOVA. Otherwise, color denotes the median difference from the expected number of transitions across simulations, with red indicating that the observed number of state switches was greater than expected, and blue indicating fewer state switches than expected. Asterisks indicate that in 90-95% (*) or >95% (**) of simulations, the number of transitions was greater (red arrows) or less (blue arrows) than the calculated expected number of transitions for the simulation in question. The percent listed under each pair of states is the median percent of the tree that is in that state-combination across all simulations (and thus do not necessarily add to 100). Panels C and D show the numbers of transitions between female song states and cooperative breeding states that occurred in lineages with C) weak territoriality and D) strong territoriality. The legend in (B) applies, qualitatively, to panels C and D.

directionality of that link could not be resolved—we found evidence to support both female song positively influencing cooperative breeding and cooperative breeding positively influencing female song, suggesting the potential for a bidirectional coevolutionary relationship (Fig. 4A). This link was robust to data availability bias corrections (Fig. 4B-D, Tables S13, S14), and alternative cooperative breeding classification schemes (Fig.

ED6). When using year-round territoriality instead of strength of territoriality classification, only the positive directional effect of cooperative breeding on female song remained robust (Fig. ED7).

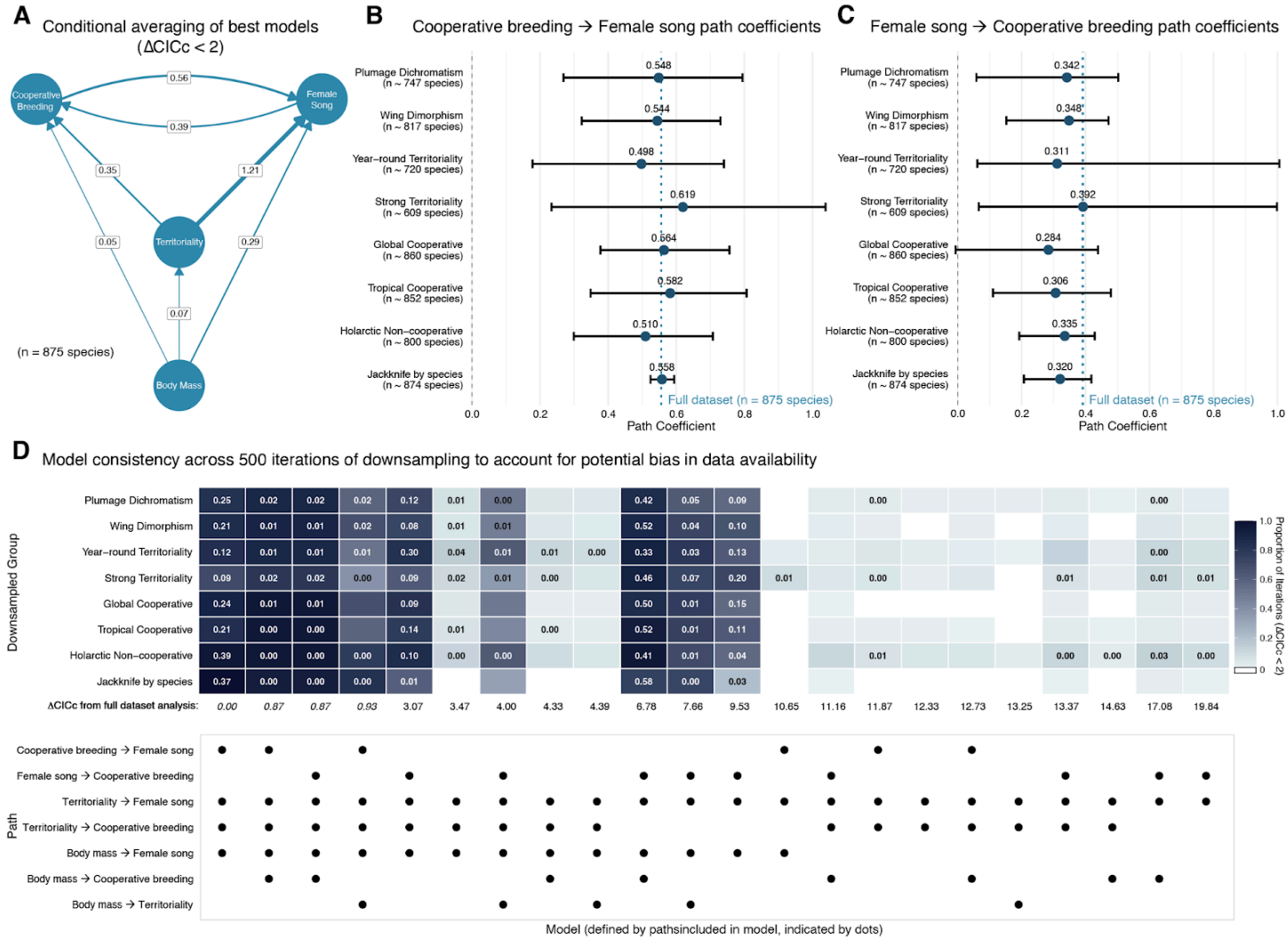


Fig. 4: The links between cooperative breeding and female song in phylogenetic path analyses are robust to correcting for data availability biases. (A) Directed acyclic graph showing conditional-averaged path coefficients from phylogenetic path analysis of the relationships between female song presence, cooperative breeding, strength of territoriality, and body mass (log-scaled and normalized). Arrow thickness indicates the magnitude of path coefficients, with all paths shown representing positive relationships. Only models within ΔCICc < 2 of the best model (4 models) were included in the averaging. (B) Forest plot showing the mean and 95% confidence interval for the cooperative breeding → female song path coefficient across 500 iterations for each bias correction type, compared to the value from the full dataset (dashed line). (C) Forest plot showing the mean and 95% confidence interval for the female song → cooperative breeding path coefficient across 500 iterations for each bias correction type, compared to the value from the full dataset. (D) Model consistency across bias corrections. Heatmap showing the number of iterations for which each phylogenetic path model was significant (had ΔCICc < 2) across 500 iterations for each bias correction type. The row underneath the heatmap contains the ΔCICc values for each model in the phylopath analysis performed using the full dataset, with the ΔCICc values of the four models that were statistically equivalent (ΔCICc < 2) in the full dataset in italics. The lower panel shows the models based on the presence or absence of each rate in the model. The models shown are all those that were in the top 20 models (out of 47 total models, see Fig. S1) by # of iterations with ΔCICc < 2 for any of the bias correction types.

To estimate the size of the effect of cooperative breeding on female song, and vice versa, in the context of territoriality and body size, we performed phylogenetic generalized linear model analyses. In the best-supported model predicting female song presence, cooperative breeding increased the odds of female song by 5.70-fold (95% Confidence Interval (CI): 2.17-9.28, p = 0.004, Table 1) in species with weak or absent territoriality, a trend that was consistent across alternative cooperative breeding classifications (Table S15). However, a significant negative interaction term (Odds Ratio (OR) = 0.24, CI: 0.16-0.67, p = 0.024, Table 1) indicated that this effect was substantially reduced in strongly territorial species. When examining the inverse relationship, female song presence increased the odds of cooperative breeding by 6.88-fold (CI: 3.61-11.28, p < 0.002, Table 1, Table S16) in weakly territorial species, but this effect was also much less pronounced in strongly territorial species (interaction OR = 0.15, CI: 0.09-0.30, p < 0.002, Table 1, Table S1). When using year-round territoriality as an alternative binary territoriality classification, the best-supported models included the associations between cooperative breeding and female song, though the estimated effect sizes were considerably weaker (CB predicting FS: OR = 1.41; FS predicting CB: OR = 1.44; Table S17) and did not include interaction terms between cooperative breeding/female song and territoriality. Importantly, the classification based on whether territories are held year-round may conflate species with fundamentally different territorial strategies—for example, by grouping strongly territorial seasonal breeders with non-territorial and weakly territorial species.

Table 1 | Best base phylogenetic generalized linear models (phyloglm in R package 'phylolm') including only territoriality, body mass, and female song or cooperative breeding as predictors. Analyses were performed for 500 bootstrap iterations. *p*-values are determined empirically based on the number of bootstrap iterations that fall above or below an odds ratio value of 1.0.

Predictor (of Female Song)	*Odds Ratio*	*CI*	*p*	*Predictor (of Cooperative Breeding)*	*Odds Ratio*	*CI*	*p*
(Intercept)	0.60	0.36-0.94	**0.028**	(Intercept)	0.12	0.05-0.26	**<0.002**
Cooperative Breeding	5.70	2.17-9.28	**0.004**	Female Song	6.88	3.61-11.28	**<0.002**
Territoriality (Weak vs Strong)	3.82	2.48-6.59	**<0.002**	Territoriality (Weak vs Strong)	6.05	3.07-9.25	**<0.002**
log(Body Mass) (normalized)	1.45	1.14-1.76	**<0.002**	Female Song:Territoriality	0.15	0.09-0.30	**<0.002**
Cooperative Breeding:Territoriality	0.24	0.16-0.67	**0.024**				
Species	875			Species	875		
Model formula: Female Song ~ Cooperative Breeding * Territoriality (Weak vs Strong) + log(Mass)				Model formula: Cooperative Breeding ~ Female Song * Territoriality (Weak vs Strong)			

Alternative social predictors of female song

To determine whether any social variables could better explain female song patterns than cooperative breeding, we used phylogenetic generalized linear models to compare each social variable against cooperative breeding while controlling for territoriality and body mass. With the caveat that our sample sizes are two- to four-fold smaller than in Table 1 because we had to exclude species missing data for either variable, we found that cooperative breeding had a higher odds ratio than all social variables except the average number of caretakers of young; in other words, when we include information on the number of individuals involved in cooperative caretaking instead of just the existence of cooperative caretaking, the model's ability to predict female song presence is even stronger (Table 2, Table S18). Including traits related to social stability, namely longer social bonds, familial living, and smaller group size, produced models that are statistically equivalent (|ΔAIC| < 2) to the models with cooperative breeding whereas including colonial living made it significantly worse (Table 2).

Table 2 | Phylogenetic generalized linear model comparisons of cooperative breeding versus alternative social predictors of female song. We compared cooperative breeding against each social variable using identical model structures with additive predictors (Female Song ~ [Social Variable] + Territoriality (Weak vs Strong) + Body Mass). Sample sizes represent species with complete data for both cooperative breeding and the social variable being compared. All ordinal or continuous social variables were standardized (scaled to have a standard deviation of 1 and mean of 0). The odds ratio for each social variable predictor is reported as the mean and 95% CI across 500 bootstrap iterations. The *p*-value is determined empirically based on the fraction of bootstrap iterations where the predictor's odds ratio falls above or below 1.0, and significant *p*-values ($p<0.05$) are bolded. AIC values that are >2 ΔAIC lower than the alternative model are highlighted in yellow. The higher of the two models' odds ratios are highlighted in green.

		Results from model using **cooperative breeding** as the predictor (Female Song ~ Cooperative Breeding + Territoriality + Body Mass)				Results from model using the **social variable** as the predictor (Female Song ~ [Social Variable] + Territoriality + Body Mass)			
Social Variable	Number of Species	Cooperative Breeding Model AIC	Cooperative Breeding Odds Ratio	Cooperative Breeding Odds Ratio 95% CI	Cooperative Breeding *p*-value	Social Variable Model AIC	Social Variable Odds Ratio	Social Variable Odds Ratio 95% CI	Social Variable *p*-value
caretakers (normalized)	242	251.8	4.65	1.57-15.29	**0.008**	**241.1**	**8.33**	2.16-31.09	**0.032**
Familial living	407	445.8	**1.91**	0.93-4.03	0.072	445.9	1.35	0.79-2.47	0.284
Colonial living	242	**251.8**	**4.60**	1.34-13.98	**0.008**	257.5	1.47	0.64-3.38	0.356
Grouping (normalized)	231	238.6	**8.37**	2.34-38.60	**0**	240.1	0.68	0.48-0.93	**0.02**
Social bond duration (normalized)	242	251.8	**4.57**	1.28-14.71	**0.016**	252.0	1.54	1.18-2.09	**<0.002**

Discussion

Our phylogenetic comparative analyses across the songbird lineage reveal a significant, bidirectional association between cooperative breeding and the presence of female song that cannot be explained by their shared associations with territoriality, body mass, or geographic sampling biases and is robust across various analytical methodologies, alternative phylogenies, and alternative methods for cooperative-breeding classification. While both traits are indeed more common in territorial species, phylogenetic path analyses and generalized linear models demonstrate an additional evolutionary link between cooperation and female song. Further, we find that this relationship is modulated by territorial context; in strongly territorial species, cooperative breeding and female song are both common and co-occur slightly more often than expected by chance, whereas in species with weak or no territorial defense behavior, cooperative breeding and female song are overall less common but co-occur much more often than expected. These dynamics suggest that female song evolution could follow distinct pathways depending on social context: whereas territorial species may maintain female song for resource defense, non-territorial cooperative breeders may have evolved female song through different selective mechanisms, potentially related to the unique social dynamics of cooperative groups. In light of these findings, we consider evidence for and against three hypothesized functions of female song in cooperative breeding systems: sexual competition between females, social conflict, and social cohesion.

The sexual competition hypothesis predicts that female song in cooperative breeding systems is promoted by intensified intrasexual competition among females due to reproductive skew[29,30]. Under this hypothesis, competition for limited breeding positions, with unsuccessful females serving in helper roles, would amplify sexual selection on females. If song mediates this competition, we would expect female song to co-occur with cooperative breeding. Female song has indeed been linked to mate defense against other females, particularly in polygynous contexts[31,32], suggesting that female song can function in intrasexual competition. Additionally, our analyses show that the evolution of male song repertoire size was slower in cooperative breeding lineages, contrasting with the accelerated song repertoire evolution seen in polygynous systems[33]. This potential evidence for reduced sexual selection pressure on males in cooperative systems

could theoretically correspond to increased pressures on females if cooperative breeding reverses typical sex differences in reproductive variance (per Bateman's principle)[15]. However, empirical evidence from cooperative breeding songbirds suggests that helping is male-biased or mixed-sex but not female-biased[34] and female helpers were not typically linked to extreme reproductive skew[13]. Together, these patterns suggest that the conditions necessary for intensified female sexual competition—female-biased helping driven by reproductive monopolization—are not characteristic of cooperative breeding songbirds, casting doubt on the hypothesis that female song evolves in these systems due to reproductive skew.

The social competition hypothesis proposes that female song mediates competition for ecological resources, particularly territories[35]. Under this hypothesis, song functions in resource defense and territorial disputes between individuals or groups, with these agonistic interactions driving the evolution and maintenance of female song. Territory defense represents one of the best-supported functions of female song across species, with females in several species using song similarly to males for solo, paired, or group defense of a resource (reviewed in [6]), and has been previously shown to be strongly associated with female song incidence[24]. Indeed, we found that strong territoriality and year-round territoriality were both significant predictors of female song presence. However, our phylogenetic path analyses showed that the cooperative breeding-female song association persists even when controlling for territoriality, and our phylogenetic generalized linear models demonstrated that the bidirectional effects linking female song and cooperative breeding are strongest in weakly territorial species. Further, if social competition for territories or other resources was the primary force driving female song evolution, we would expect female song to be similarly associated with other high-density social systems that presumably intensify resource competition. However, we found that cooperative breeding was a stronger predictor of female song than colonial living and familial living for both methods of territory classification and that larger group sizes actually predicted an *absence* of female song. Together, these patterns suggest that, while social competition may drive the presence of female song in some contexts, it does not account for the specific association with cooperative breeding, especially in species with weak or absent territorial behavior.

A third hypothesis is that female song could serve a role in maintaining social cohesion in cooperative breeding groups. Song produced in the context of established breeding pairs is already well-documented as a

factor in reinforcing pair bonds; however, evidence suggests that song can also mediate affiliative social interactions beyond the breeding pair, potentially by co-opting the reward systems that support pair-bond maintenance. For instance, European starlings' "gregarious song", which occurs outside of the breeding season, is positively correlated with social flocking behavior and is thought to be intrinsically rewarding[36,37]. Male zebra finch "undirected song" (i.e. not directed toward a female) occurs year-round and is hypothesized to have prosocial functions, such as breeding synchronization, promoting social cohesion, and individual recognition[38]. Female song shares some patterns with these examples (reviewed in[4]): it often occurs both in and out of the breeding season[39] and has been linked to breeding-related contexts including coordination of breeding activities[4,40] and interactions with offspring post-fledging[41], and non-breeding-related functions including helper retention[42], individual or group-member recognition[42,43], neighbor recognition[44], and social cohesion[45]. These affiliative functions suggest that song could strengthen cooperative interactions through group cohesion, bonding, and coordination in species where cooperation and altruism are key life history strategies. While social living may often lead to more frequent competition among females for mates or other resources, the maintenance of the social group—and the fitness benefits thereof—should select for behaviors that minimize sexual conflict, particularly in systems with both high relatedness and significant cooperation. While social bond duration and the number of caretakers also showed positive associations with female song, most other social variables showed weak or negative associations. This pattern, where cooperative caretaking and longer social bonds promote female song, but other forms of group living do not, supports the hypothesis that female song may serve prosocial functions specifically in stable, cooperative social systems. Further, species where female song already mediates group coordination or offspring communication may be pre-adapted for transitions to cooperative breeding, as the vocal behaviors necessary for managing complex multi-caregiver systems would already be in place.

The sex-specific patterns in our data provide additional support for distinct evolutionary pressures on male and female song in cooperative species. Male song repertoire size evolved more slowly not only in cooperative breeding lineages, but also in familial lineages, regardless of whether cooperative breeding was present. This pattern suggests that familial living and cooperative breeding can each stabilize male song elaboration, while only kin-cooperative breeding promotes or stabilizes female song. These contrasting results

may hint at sex-specific functions and selection pressures underlying male and female song evolution. For example, our results could suggest that female song, but not male song, has group-cohesive functions, since female song is most promoted in the system in which cohesion should have the greatest fitness benefit, i.e. kin-cooperative systems[34]. Our results might also suggest that male song, moreso than female song, is important for kin recognition and provides a mechanism for inbreeding avoidance, in which case male song repertoire size might be conserved in systems in which extended kin contact leads to a higher risk of inbreeding, i.e. both kin-cooperative breeding and familial living[46].

Our findings regarding song repertoire evolution alongside other social factors may provide additional context. In addition to non-cooperative, non-familial systems, song repertoire size appears to evolve more rapidly in pair-based or large-group social systems, lineages with shorter social bonds, and colonial living systems, in addition to polygynous mating systems (identified in [33]). All together, these results suggest a mixed-selection-pressure paradigm in which male repertoire sizes evolve faster in environments where there is greater competition for mates due to increased social density and less social cooperation, and are conserved when there is reduced competition among males for mates, shared effort among groupmates in defending territory, or pressure against becoming too elaborate and disrupting cohesion within a stable group.

Exposure to challenging or variable environmental conditions may provide another explanation for the link between cooperative breeding and male song repertoire stability, as complex songs are hypothesized to be disproportionately affected by stressful environmental conditions during development[47], and such conditions are also theorized to contribute to the emergence of cooperative breeding[48] or defensive behaviors[49]. The latter theory is bolstered by findings suggesting that cooperative breeding is more common in bird species found in environments with highly variable climates such as Sub-Saharan Africa, Australia, and the tropics[50–52]. Several studies have also linked cooperative breeding with life history traits indicative of individual robustness; for instance, linking cooperative breeding with longer post-fledgling feeding periods and longer lifespans[53,54], both of which are expected to be linked with overall less stressful development. These trends could have implications for song evolution since developmental conditions can have significant effects on young birds; a stressful upbringing can lead to impairments in both song learning and the development of the neural song system, while extended parental provisioning appears to promote the evolution of larger brain sizes across

species [25,55]. Extended parental care and cooperative breeding could thus help maintain song-learning capacity in harsh or highly variable environments, potentially stabilizing the song over time. Indeed, consistent with this hypothesis, we find slower song repertoire evolution in both cooperative breeding lineages and lineages with multi-year social bonds.

Our findings could be clarified by recategorizing species with female song based on the level of sexual dimorphism in song features or production or specific contexts of female song, such as duetting. In the longer term, it would be highly informative to have species-level characterizations of female song in addition to male songs. Finally, it has been suggested that cooperative breeding should be considered a three-state trait, distinguishing species that exhibit occasional cooperative breeding from those in which it is either common or absent[56,57]; separating the cooperative breeding trait as such might reveal more complex trends. Additionally, expanding datasets for other social variables beyond cooperative breeding would enable more comprehensive analyses of how different forms of sociality influence vocal evolution.

In conclusion, we find robust evidence for context-dependent associations between cooperative breeding and singing behavior across the songbird phylogeny, with territorial systems mediating when and how these traits coevolve. While cooperative breeding and female song co-occur at high rates in strongly territorial systems, their mutual reinforcement is especially pronounced in weakly territorial systems. Our findings also suggest that the context-dependent relationship between cooperative breeding and female song is best captured by territorial classifications that reflect how vigorously species defend territories rather than simply the duration that the territory is held. The difference in the relative effect of cooperative breeding in different territorial systems suggests that female song evolution follows multiple pathways: territorial defense in competitive environments and social coordination in cooperative contexts. While male song repertoire size evolution is constrained by various forms of stable social organization, female song shows robust associations with cooperative breeding that persist even when controlling for territoriality, body size, or other life history factors. This relationship, combined with its context-dependent nature, suggests that female song can serve important functions in facilitating the coordination required in cooperative breeding systems. Our findings underscore the idea that song can have social functions beyond its canonical roles in mate choice and territory

defense and reveal how cooperative living can shape communication systems in fundamentally different ways depending on the ecological and social context.

Methods

Compiling the database

Cooperative Breeding: We found previously published data on cooperative breeding using the specific search criteria “Cooperative breeding in birds” and “Songbird cooperative breeding” on Google Scholar and Web of Science, yielding 11 sources with previously compiled species information on cooperative breeding[10,16,34,50–53,56,58–60]. We considered cooperative breeding behavior to be alloparenting during incubation, brooding, and nest provisioning in 10% or more of nests in a study, if quantitative data were available, following the classification methods of Cockburn[10]. We did not consider nest building, territory defense, or fledgling provisioning by extrapair individuals to be sufficient for cooperative breeding classification. When a source’s data were qualitative, species described unambiguously as “cooperative,” “breed cooperatively,” “often cooperative,” “facultatively cooperative breeding,” or as commonly having “helpers” at the nest were considered to be “cooperative”, and species described with “(cooperative breeding) does not occur,” “none,” “extremely rare” were considered to be “noncooperative.” We did not consider “Not observed,” “Not known to occur,” or “Never reported” to be sufficient evidence for the absence of cooperative breeding. If a source described a species as having “occasional,” “possible,” “suspected,” “inferred,” “sometimes,” or “infrequent” cooperative breeding, we did not assign a binary classification to the species from that source[56]. Cooperative breeding classifications were assigned to each species using a binary system based on the aggregation of classifications from unique sources describing the species as cooperative or non-cooperative; when one source cited another source included in our dataset we did not consider the citing source to be unique. Species for which source classifications were not unanimous were, when possible, classified based on the description in Birds of the World, otherwise they were omitted from our primary analyses (HighConfidence_Coop). To ensure that our results were robust, we repeated any analyses that produced significant results using alternative classification methods based on whether the majority of unique sources classified a species as

cooperative versus non-cooperative. For those species that were “tied” with equal numbers of sources classifying them as non-cooperative and cooperative, we ran secondary analyses with these species 1) classified as non-cooperative (MeanCoopTie2Noncoop), 2) classified as cooperative (MeanCoopTie2Coop), and 3) omitted from the dataset (MeanCoopOmitTies). We also ran analyses with 4) any species with at least one source classification of cooperative breeding designated as cooperative (AnyCoopEqualsCoop) and 5) any species with at least one source classification of non-cooperative breeding designated non-cooperative (AnyNoncoopEqualsNoncoop) (see Table S1). We also repeated analyses using the subset of classifications from several individual sources (see S2 Dataset). We also tested the effect of recategorizing cooperative breeding based on the “Inferred” parental care mode, rather than the higher-confidence “Known” classifications from [10].

Other life history metrics: To tease apart the effects of breeding systems and broader social structures, we also obtained species classifications for several metrics of sociality from [25]. We binarized several categorical metrics from that study related to group size, duration of social bonds, colonial living, and number of caretakers (Table S1). We also included binary classifications of species as socially monogamous or socially polygynous from [33], which classified social polygyny primarily based on having an observed rate of polygyny above 10% in studied breeding pairs [33], and as having familial or non-familial living from [16], which classified familial living as species where offspring remain with parents for at least 50 days post-nutritional independence. We also included multi-state traits for social group size from [25] and the interaction of familial living and cooperative breeding from [16]. In our phylopath and phyloglm models, we incorporated log(Body Mass) from [61].

Territory: We obtained Territory data from [21], which categorizes territorial behavior into three states: non-territorial (level 1), seasonal or weak territoriality (level 2), and year-round territorial (level 3). However, the second level of this categorization groups species with strong seasonal territoriality together with species that are weakly territorial, which might lump species with very different territorial strategies while also separating species that exhibit strong territorial defense in temperate regions from those with year-round territory defense primarily in tropical regions. To better understand the relationship between our focal traits and the intensity of territoriality, we refined this classification into "weak" versus "strong" territorial defense categories.

Non-territorial species (Territory level 1) were assigned to the "weak" category and year-round territorial species (Territory level 3) to "strong"; species with seasonal or weak territoriality (Territory level 2) were individually recategorized by incorporating additional information from published sources. For the 523 species classified as Territory level 2 for which we had both female song and cooperative breeding data, we reviewed species accounts in Birds of the World[62] supplemented by Google Scholar literature searches when necessary. To ensure reproducibility, 30 species were independently classified by a second researcher. Our classification criteria for **strong territoriality** included: clear territorial defense behaviors (e.g., "vigorously defending," "chasing intruders," "boundary disputes"), well-defined territory boundaries with minimal overlap, persistent defense throughout breeding season or year-round, active defense mechanisms including displays and vocalizations, and specific territorial use of song. Key indicators included descriptions such as "highly territorial," "aggressively defends," "non-overlapping territories," "responds aggressively to playback," and "maintains territory." **Weak territoriality** was characterized by: limited spatial defense (small areas around nest sites only), tolerance of conspecifics within defended areas, significantly overlapping territories, colonial or semi-colonial breeding, minimal defense behaviors, and mate-focused rather than territory-focused defense. Indicators included "loosely colonial," "overlapping boundaries," "defends only immediate nest site," "soft territoriality," and "tolerant of intruders." This reclassification resulted in 170 species classified as weak territorial (including 114 species with Territory level 1 plus 56 reclassified species with Territory level 2) and 705 species as strong territorial (including 404 species with Territory level 3 plus 301 reclassified species with Territory level 2), with 166 species with Territory level 2 remaining unclassified due to insufficient information.

Female Song: In addition, we combined and updated published databases on the presence of female song in songbird species in natural conditions[2,63–65], S3 Dataset), which we located by performing searches on Google Scholar and Web of Science for “Female song passerine/passeriformes” and “Female song songbird”. We considered a species to have female song if it was classified as such in one of the published female song databases or if a source unambiguously said that females sing. Whenever sources disagreed on female song classification, we first attempted to resolve these discrepancies by cross-referencing with the species entry in Birds of the World[62]. If one of the sources for female song in a species with source discrepancies included Garamszegi *et al.*[64], we also cross-referenced these classifications with their original source in the

encyclopedia *Birds of the Western Palearctic*[66]. We also used data from Mikula *et al.*[67], but since the species classifications in that source created binary variables for female solo song, absence of female song, and duetting (a category that included chorusing species, which may include species in which only males sing), we only used species classifications that indicated that females produced solo song or did not sing. We further checked Birds of the World if a species was very well studied (i.e. had at least one publication noting species song features and at least one publication regarding breeding behavior in the species) but did not appear in any female song database. Species in which females had only ever been documented singing after treatment with estradiol or testosterone were omitted from our dataset, even if another source classified them as having female song present. Documentation of subsong in females in a species was not considered sufficient for a classification of female song present. We were ultimately able to assign species female song classifications for 1094 species. While the diversity in female singing behaviors, song complexity, and their difference from male songs are crucial areas for future research, for the purpose of maintaining sufficient sample sizes with the data available, we opted not to assign female song classifications along more granular categories related to song complexity or song difference from males.

Song features: We obtained species data on syllable repertoire size (average number of syllables a bird can produce), syllables per song (average number of unique syllables a bird sings per song), song repertoire size (the number of individual songs a bird knows), song duration (the length of a song), and inter-song interval (time between songs) from Snyder and Creanza[33] and added several more species-level data points to the database (S1 Dataset). The primary sources of the song-feature data were generally field studies of presumably male birds; when song-feature data were specified by sex in the literature, we used only the male song measurements. Whenever we obtained data for a species whose name was not present in the BirdTree.org phylogeny, we searched Avibase for alternative taxonomic classifications. If an alternative species name was present in the BirdTree phylogeny (e.g. the genus or species name has been updated), we recorded the species data in our database under the BirdTree-matching name.

Dimorphism: We quantified both sexual dimorphism (body size) and sexual dichromatism (plumage coloration). For plumage dichromatism, we used species- and sex-specific scores from [60], calculating the absolute difference between male and female plumage scores and applying a log transformation with an offset

527 of 0.1 to handle zero values. We were able to obtain a plumage dichromatism value for all of our 1041 species
528 with both cooperative breeding and female song data. For size dimorphism, we used morphometric data from
529 the AVONET database[61], focusing on wing length as a proxy for body size. We log-transformed raw wing
530 measurements, calculated sex-specific means for each species, and quantified dimorphism as the percent
531 absolute difference between male and female log-transformed wing lengths relative to female values. This
532 percentage metric accounts for allometric scaling and provides a standardized measure of size dimorphism
533 across species with different body sizes. We were able to obtain wing dimorphism values for 987 of our 1041
534 species with cooperative breeding and female song data.
535
536 Generating a consensus phylogeny
537 To analyze evolutionary patterns in light of the relationships between species, we downloaded 1000
538 trees randomly sampled from a pseudo-posterior distribution of trees using the Hackett backbone from
539 BirdTree.org[27,68]. We subsetted these trees to include only the 5,966 passerine species (including Oscines and
540 Suboscines) represented in the trees. We then used the consensus.edges function (R package: phytools, with
541 method="mean.edge" and the default setting if.absent="zero", which treats missing edges as zero in the mean
542 edge calculation, but see "Accounting for phylogenetic uncertainty" below for other approaches) to build a
543 consensus tree with branch lengths that included all passerine species. For all subsequent analyses, we
544 removed the Suboscine clade and only Oscine species were included in reference phylogenies.
545
546 Estimating rates of transition between states of categorical traits
547 We conducted ancestral character estimations (function ace, R package: ape) using the maximum
548 likelihood method to estimate the evolutionary history of discrete traits (e.g. cooperative breeding, sociality
549 metrics, female song presence) across all Oscine species for which we had a classification for a given trait.
550 Using a likelihood ratio test, we compared the fits of the ancestral character estimation model that assumed
551 "Equal Rates" (ER) of transition between binary ("0" and "1") states and the model that assumed "All Rates
552 Different" (ARD, i.e. that 0→1 and 1→0 transitions occur at different rates) (Table S19). For traits with three or
553 more states, we also tested the model that assumed "Symmetric" (SYM) rates. In subsequent analyses, we

554 used the ARD model whenever it was significantly better than the ER model (and SYM model, when
555 applicable). Whenever we generated models of trait evolution on subsets of species, to avoid any potential
556 bias from recalculating rates for each species subset, we used the transition rates between states that were
557 obtained from the full set of Oscine species that had a state classification for that trait.
558
559 Song feature evolution in the context of cooperative breeding and other categorical traits
560 To determine whether there was a difference in the values of various song characteristics between
561 groups of species that exhibit certain social traits, e.g. cooperative and non-cooperative breeding species, we
562 used a set of phylANOVA analyses (R package: phytools,[69]) to compare natural log-transformed song
563 characteristics (obtained from Snyder and Creanza[33]) between the sets of species in the different social trait
564 categories, with 50,000 simulations per song feature + social trait comparison[33].
565 We then tested whether any song characteristics evolved at different rates when in a cooperative
566 breeding state versus a noncooperative breeding state using the Brownie algorithm[26]. We generated 500
567 simulated stochastic character maps ("simmaps"), which simulate ancestral gains and losses of cooperative
568 breeding given the transition rates calculated above, using the function make.simmap (R package: phytools),
569 and used them in the brownie.lite function (R package: phytools). We performed a follow-up Brownie analysis
570 using each of the binarized sociality traits and song repertoire size, which was found to consistently evolve
571 faster in noncooperative than cooperative lineages across simulations. We further performed Brownie analyses
572 on song repertoire size evolution and the multistate traits of group size ("grouping" from [25]), the intersection of
573 familial living and cooperative breeding traits from [16], and the intersection of our cooperative breeding and
574 territoriality strength classifications, with transition rates obtained from ape::ace() with the All Rates Different
575 (ARD) setting, after testing and confirming that the ARD model was significantly better than both the Equal
576 Rates (ER) and Symmetric Rates (SYM) settings.
577 Species for which there were multiple values for a certain song feature in the literature have their
578 minimum and maximum values recorded in the feature dataset. To account for potential variations in song
579 feature measurement across species, researchers, locations, and decades of study, for the feature that
580 showed significant dependent evolution (cooperative breeding with song repertoire size), we repeated the

581 Brownie analysis with only the minimum values, with only the maximum values, and with random sampling
582 from the minimum, median, and maximum values for all species, where we resampled the song repertoire
583 values 500 times, generating 2 simmaps per resampled vector to conduct the analyses of evolutionary rates.
584
585 Analyzing evolutionary co-occurrence of cooperative breeding, female song, territoriality, and other traits
586 To test whether pairs of discrete traits—including cooperative breeding, other sociality classifications
587 (Table S1), and female song presence—co-occurred in the songbird lineage more often than expected by
588 chance, we employed methods adapted from [28]. This process involves generating stochastic character maps
589 (simmaps) using phytools::make.simmap() for the two discrete characters in question (for example,
590 cooperative breeding and female song) and setting the rates of transition between states (Q) to those
591 estimated using ape::ace(). Then, for each pair of simmaps we calculated the proportions of the branches of
592 the phylogeny spent in each of the four state combinations arising from combinations of the two binary traits (or
593 more than four when testing one of the multi-state traits versus female song) using phytools::map.overlap().
594 We then calculated the statistic *D* (equations given in [28]) to obtain an overall measure of the disagreement
595 between observed and expected proportions of each combination of characters. Per [28], we repeated this
596 process to generate simulations of the two discrete traits evolving independently from one another. To do this,
597 we randomly shuffled the tip states at the same proportion as the actual species classification data, then
598 generated each simmap with phytools::make.simmap(). The empirical *p*-value is calculated as the fraction of
599 randomized-data simmap pairs that have a *D* statistic larger than the median *D* statistic of the simmap pairs
600 generated from the real data. We also tried using phytools::sim.history() to generate the simulated independent
601 evolution simmaps using the rates of transition between noncooperative breeding and cooperative breeding
602 and between female song absent and present obtained from make.simmap(). We found that these rate-based
603 stochastic character maps that were simulated forward in time from the root of the tree toward the tips, rather
604 than stochastic character maps that were simulated backward in time from the tips toward the root, resulted in
605 very different proportions of the tree falling in each state, and that the dominant state was determined primarily
606 by which state was randomly chosen as the ancestral state in sim.history(). Therefore, we opted to use the set
607 of simulated-independent simmaps for each trait that were made using randomly sampled tip states and

make.simmap(). Unless otherwise noted, we performed 500 real and 500 simulated-independent simulations per trait comparison. We also note that the methods of [28] can be misleading when one or more of the studied traits are clustered in one small section of the tree or have one major origin in evolutionary history [70]; however, when we visualize these traits on a large phylogeny, we observe many gains and losses of both female song and cooperative breeding throughout the tree (Fig. 2).

<u>Testing evolutionary associations and interactions using simulated state transitions</u>

Additionally, we used the simmaps generated using the real tip states to estimate whether the state of one trait influenced the likelihood that certain transitions occurred in the other trait. For each pair of trait simmaps, denoted here as trait A and trait B, we identified every state transition in each direction (trait $A_{0\rightarrow1}$ and trait $A_{1\rightarrow0}$) on the trait A tree, then identified which state was present at the same point on the trait B tree, and vice versa. This gave us 8 observed transition counts, “trait $A_{0\rightarrow1}$ in trait B_0”, “trait $A_{0\rightarrow1}$ in trait B_1”, “trait $A_{1\rightarrow0}$ in trait B_0”, “trait $A_{1\rightarrow0}$ in trait B_1”, etc. We calculated the expected number of transition counts for each simmap pair by multiplying the total number of trait A transitions in each direction by the fraction of the tree spent in each state of trait B, and vice versa. We log-transformed these state-switch counts across the 500 simulations and fit a two-way linear model with count type (observed vs. expected) and state transition type (eight possible transition types per pair of binary traits) as fixed factors, including their interaction. An ANOVA on this model tested whether the difference between observed and expected counts varied among transition categories. When the interaction term was significant, we conducted Tukey-adjusted pairwise comparisons of estimated marginal means within each transition type to identify which specific transitions differed significantly from the expected counts. To assess the effect of territoriality on the interactions between cooperative breeding and female song, we also generated a third set of simmaps for our binary weak vs. strong territoriality trait and quantified the transitions between female song and cooperative breeding states depending on the territoriality state at the corresponding point on the third simmap.

633 Testing evolutionary associations and calculating effect sizes using phylogenetic generalized linear modeling

634 We also used phylogenetic logistic regression (phyloglm() from R package: 'phylolm'; [71,72] to test for
635 evolutionary associations between traits in several ways.

636 Primarily, we used phyloglm to test for associations between female song presence and cooperative
637 breeding while controlling for territoriality (binary measures of territory strength and year-round territoriality) and
638 body mass (log-transformed, centered, and scaled) on our full dataset. We fitted multiple competing models
639 with different interaction structures and used AIC model selection to identify the best-supported trait
640 associations. For the best-fit model, we ran 500 bootstrap iterations to quantify uncertainty in the effect sizes of
641 the predictors; we report the odds ratio (OR), 95% OR confidence interval, and p-value for each predictor.

642 We integrated additional variables through two additional methods: 1) direct comparison of identical
643 models that include either cooperative breeding or an associated social variable as the main predictor,
644 alongside territoriality and body mass, in order to determine whether any social variables better predicted
645 female song than cooperative breeding, and 2) forward stepwise model expansion to assess whether any life
646 history variables improve or alter the best-fit models predicting female song and cooperative breeding. For
647 each model discussed here, we performed 500 bootstrap iterations, and we report the odds ratio, 95%
648 confidence interval, and *p*-value for the parameters in question.

649 First, to address whether other social variables might better predict female song than cooperative
650 breeding, we performed direct model comparisons. For each social variable (colonial living, familial living,
651 group size, social bond duration, number of caretakers), we compared models using that variable versus
652 cooperative breeding (alongside territoriality and body mass) as predictors, applied to the subset of species
653 that had data for both cooperative breeding and the social variable in question. Notably, since the number of
654 species for which there is data available for the various social variables is much smaller than our full
655 cooperative breeding-female song dataset, the relative performance of cooperative breeding versus other
656 social variables should be interpreted with caution, as the reduced datasets may not capture the full
657 phylogenetic or ecological diversity of our complete sample.

658 Second, we integrated additional life history variables through a forward stepwise selection process.
659 Starting with the best model from each of the four main analyses (female song or cooperative breeding as the

660 response variable; territory strength or year-round territoriality as the territoriality variable; best model
661 determined as the lowest AIC), we iteratively tested the addition of candidate predictors: absolute latitude
662 (range centroid, centered and scaled; [61]), wing dimorphism (percent absolute difference in log-transformed
663 wing length between sexes, centered and scaled; [61]), plumage dichromatism (log-transformed absolute
664 difference in male-female plumage scores, centered and scaled; [60], migration (3-level, categorical: sedentary,
665 partially migratory, migratory, treated as numeric; [61]), familial living (binary; [16]), social mating system (binary
666 social monogamy vs polygyny; [33]), and geographic region (binary: holarctic vs. tropical; [50]). At each iteration,
667 we added each remaining candidate predictor to the current model, refit both the base and the expanded
668 models using only species with complete data for all variables, and selected the predictor that yielded the
669 greatest AIC improvement. We accepted predictors that improved model fit by ΔAIC >= 2 and continued
670 iterations until no remaining predictor met this threshold.
671
672 Testing correlated evolution and causal relationships with phylogenetic path analysis
673 We used phylogenetic path analysis (R package 'phylopath'[73]) to test for correlated evolution among
674 female song presence, cooperative breeding, and territoriality, as well as body size. This method compares
675 alternative causal models while controlling for phylogenetic non-independence, allowing us to test whether
676 these traits evolved independently or show evolutionary interdependence. We specified competing models
677 representing different hypothetical causal relationships (e.g., cooperative breeding → female song vs. female
678 song → cooperative breeding vs. independent evolution) and used C-statistic information criterion corrected for
679 small sample sizes (CICc) model selection to identify the best-supported evolutionary scenario. For each set of
680 covariates, we calculated model-averaged standardized path coefficients across all models with a CICc value
681 within 2 units of the best-supported model, with coefficients weighted by relative model support. All models
682 were tested both with body mass included as a covariate to control for known and hypothesized allometric
683 effects and without body mass (full set of models provided in Fig. S1).
684
685

686 Assessment of data biases

687 To address potential systematic biases in data availability, we tested whether the species that had
688 classifications for female song and cooperative breeding varied systematically across ecological, behavioral,
689 morphological, and geographic categories. In other words, we tested whether female song and cooperative
690 breeding data were more likely to be collected in certain contexts (e.g. from species with readily distinguishable
691 males and females, from species that defend clear territories), reflecting biases in human research effort rather
692 than avian biology. We used chi-square tests to examine associations between data availability and binary
693 traits including cooperative breeding status, coloniality[25], mating system (monogamous vs. polygynous),
694 territoriality classifications, and geographic region (Holarctic vs. Tropical, based on [50]). For sexual
695 dichromatism, we calculated the absolute difference between male and female plumage scores from [60] and
696 tested for differences in dichromatism between species with and without female song data using Welch's
697 t-tests. We approximated size dimorphism by averaging by species all wing length measurements available for
698 males and females in the AVONET database[61], log-transforming the mean wing lengths by species and sex,
699 subtracting the female mean from the male mean, taking the absolute value, and dividing by the female mean
700 to get the percent dimorphism per species. We also tested whether species with female song present versus
701 absent differed in sexual dichromatism or size dimorphism, as this could indicate detection bias where sexually
702 dimorphic species are more likely to have confident female song classifications. Notably, among species with
703 female song data, those classified as having female song absent had higher dimorphism than those with
704 female song present (plumage: $t = 3.83$, $p < 0.001$; wing: $t = 4.18$, $p < 0.001$; Table S20).

705 For those traits that did show significant biases in female song or cooperative breeding data availability,
706 we performed repeated phylopath analyses, with models including cooperative breeding, female song, weak vs
707 strong territoriality, and body mass (log-transformed) by downsampling the dataset to randomly/systematically
708 remove species from groups that we found to be overrepresented. We did this for each of: holarctic
709 non-cooperative species, tropical cooperative breeding species, global cooperative breeding species, and both
710 wing and plumage dimorphism (Table S14). For the continuous-variable dimorphism metrics, we performed
711 downsampling corrections by randomly removing dimorphic species in proportion to their excess dimorphism

above the mean of species without female song data, repeating phylogenetic path analyses across 500 iterations.

Testing whether observed trends are driven by individual families

We repeated all analyses comparing song traits and cooperative breeding with significant results by performing a jackknife analysis by family, i.e. iteratively removing each family from the dataset and recalculating the results to ensure that no single family was driving the observed trends. Family classifications were obtained from AVONET[61].

Accounting for phylogenetic uncertainty

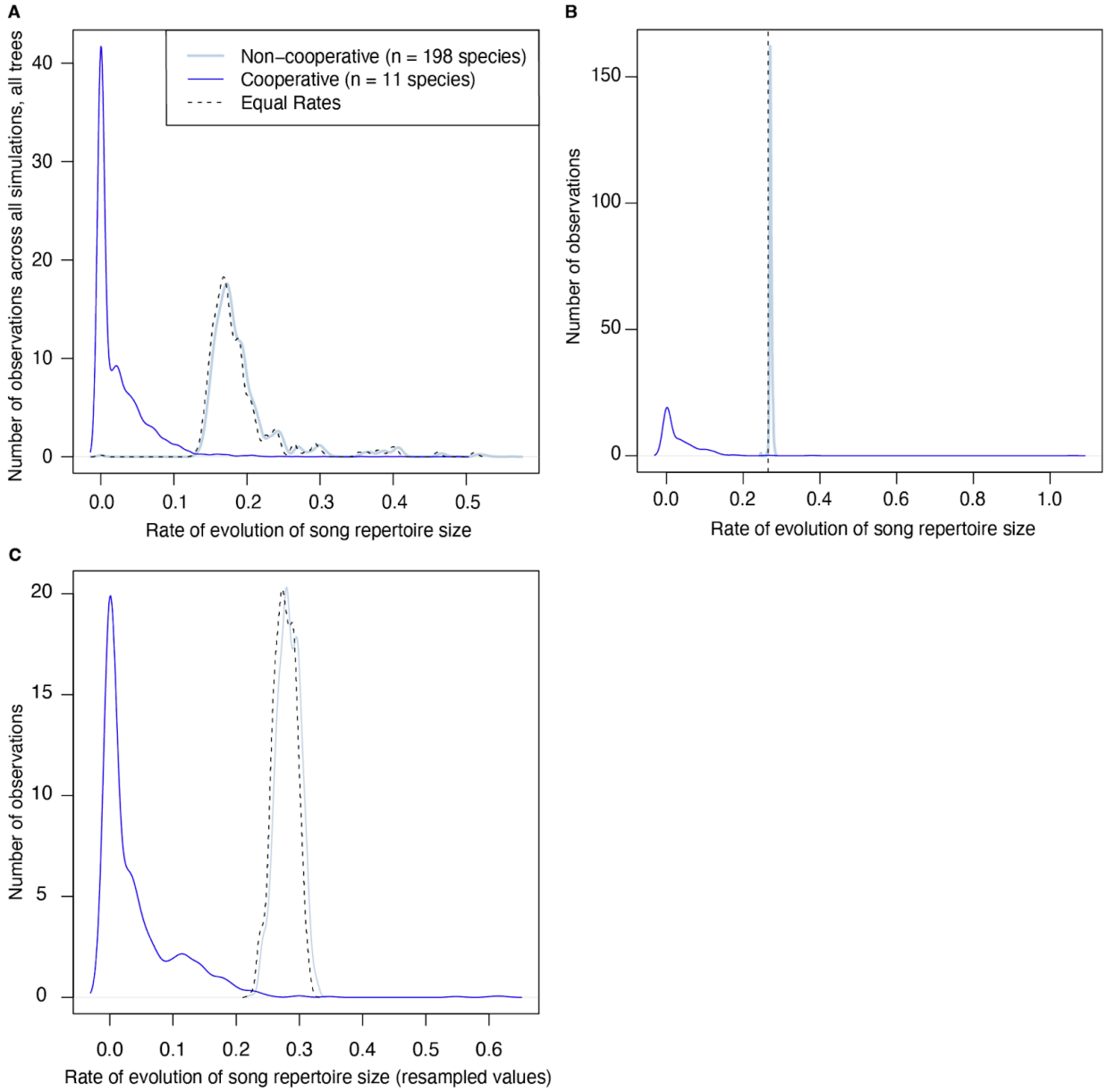
We built a second consensus tree from the same sample of 1000 trees from BirdTree.org (Hackett-Stage2, trees 3001-4000) using the consensus.edges function (R package: phytools) with method “mean.edge” but ignoring missing edges in the edge length calculations (if.absent="ignore"). To test whether this alternate method of calculating branch lengths meaningfully affected the results, we used this consensus tree to repeat tests that found significant associations between song features and cooperative breeding and between female song and cooperative breeding. To further account for phylogenetic uncertainty, we re-ran these same analyses on 200 individual trees from BirdTree.org, with 20 simulations per tree. For these, we used the transition rates calculated using the default consensus tree to generate stochastic character maps on each individual sampled tree.

Acknowledgements

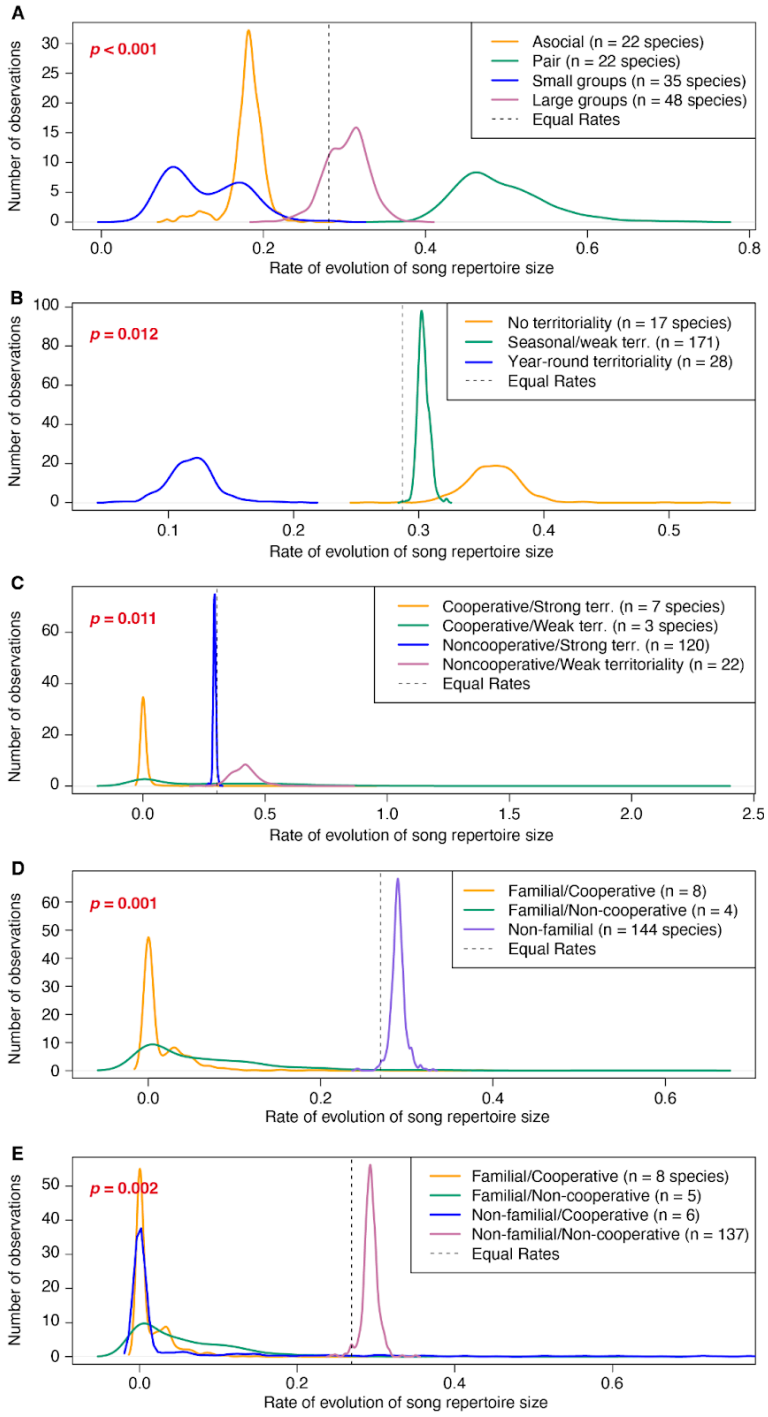
We are grateful to members of the Creanza lab for helpful feedback, and we thank the field biologists who collected all the behavioral data we compiled for the dataset. K.T.S. and N.C. were funded by grants from the National Science Foundation (IOS-2327982 and DUE-1926736). A.L.-P. was supported by the School for Science and Math at Vanderbilt. K.T.S., A.L.-P., and N.C. were supported by a Pilot Grant from the Evolutionary Studies Initiative at Vanderbilt.

Data and code availability

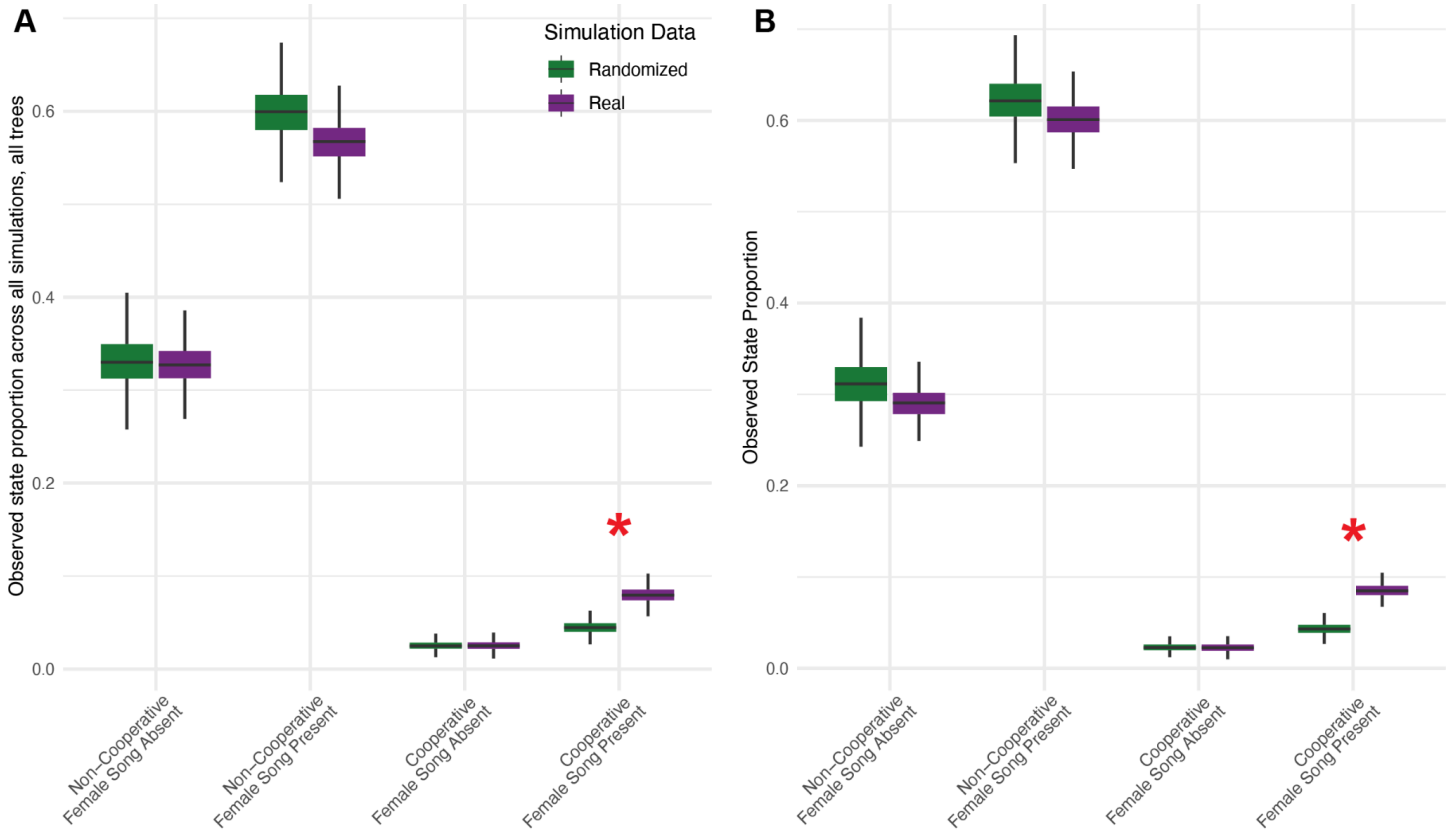
A private repository of the data and code is available at https://figshare.com/s/7e9da2429ee8e50c279d for peer review; data and code will be made publicly available before publication.



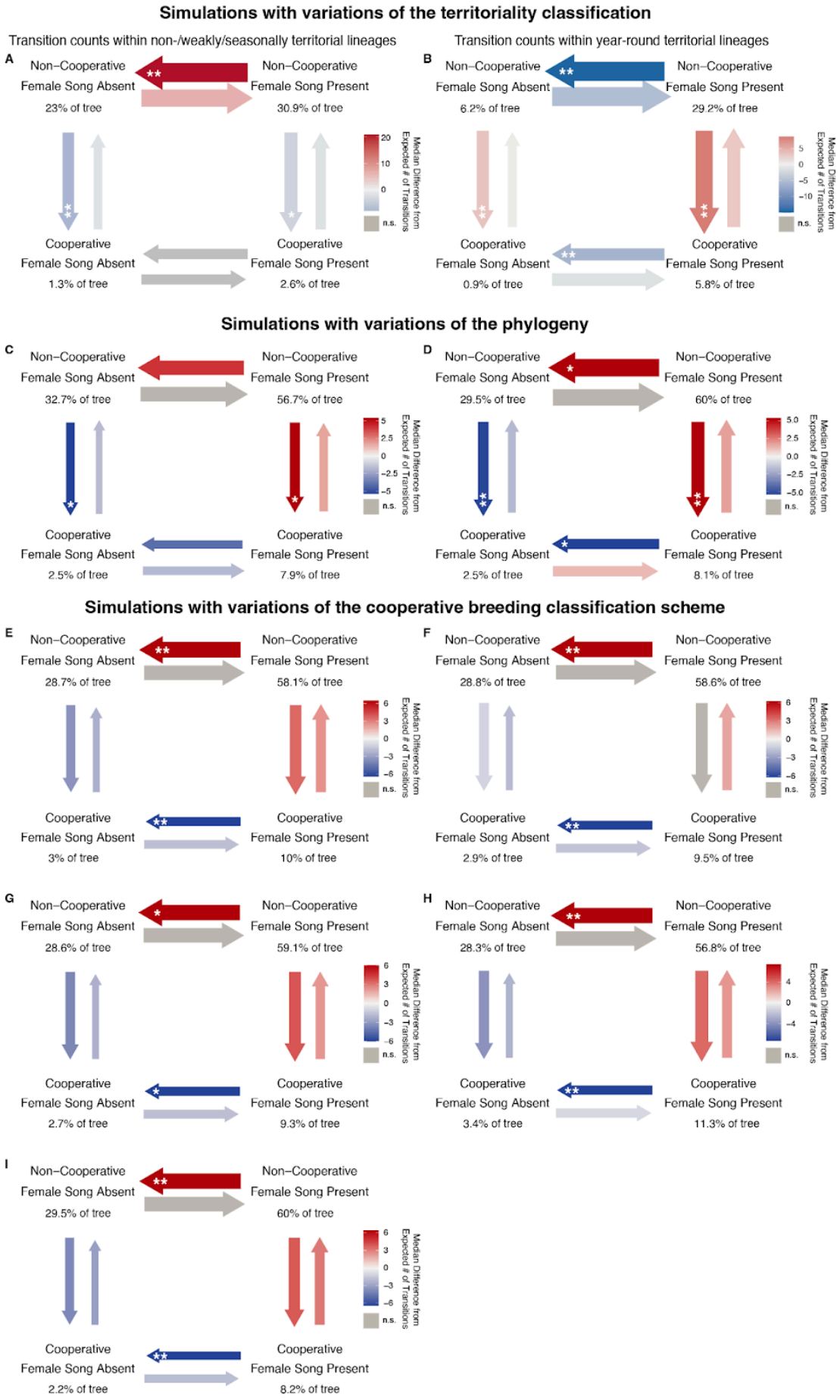
Extended Data Fig. 1: Rates of evolution of song repertoire size relative to cooperative breeding state using alternative methods to account for phylogenetic uncertainty: A) using 200 trees sampled from BirdTree.org (Hackett backbone, Stage 2), with 20 simulations per tree; B) using an alternative consensus tree (calculated with absent edges ignored in the mean edge length calculations); C) using 500 iteratively resampled minimum/median/maximum song repertoire values, with 2 simulations per resampled data vector. n = 209 species. The legend in (A) applies to all panels. See Table S4.



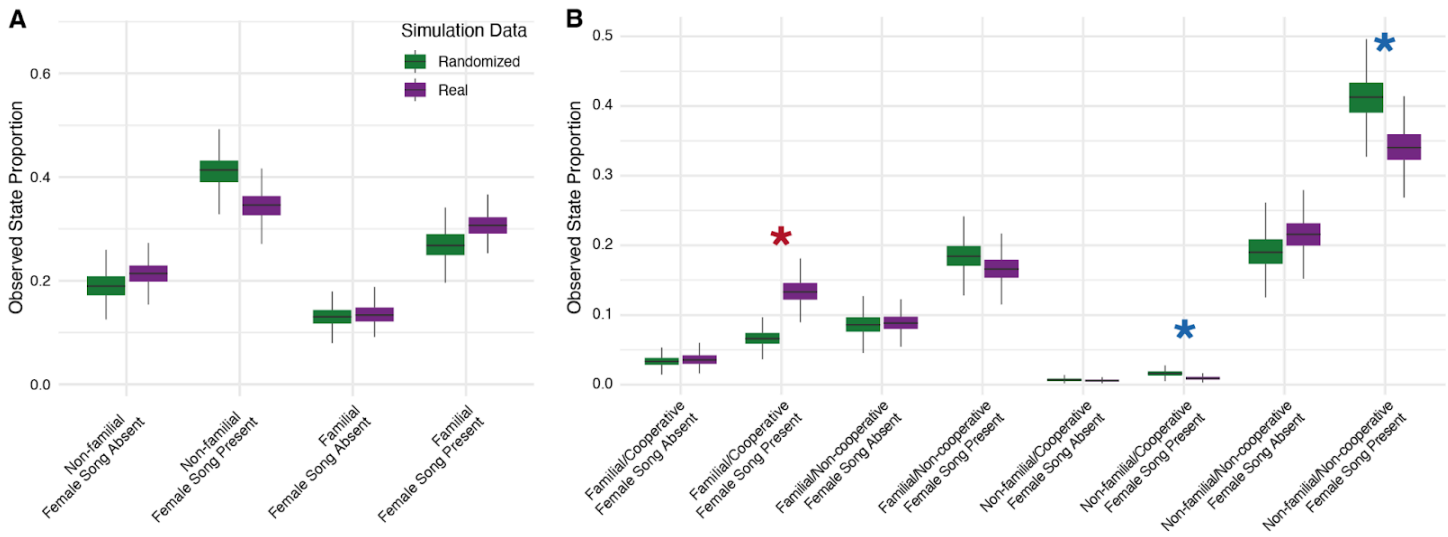
Extended Data Fig. 2: Results assessing whether the rate of evolution of song repertoire size differs based on the predicted ancestral state of (A) group size (from Griesser *et al.* 2023), (B) territoriality (from Tobias *et al.* 2016), and (C) the combination of cooperative breeding and weak vs strong territoriality, (D) a social system trait that could take three possible values, “Familial/Cooperative”, “Familial/Non-cooperative”, and “Non-familial” (“social system” from Griesser *et al.* 2017) or (E) a social system that could take four possible values, “Cooperative, Familial”, “Cooperative, Non-familial”, “Non-cooperative, Familial”, and “Non-cooperative, Non-familial” (“social system including nonkin-cooperative breeding” from Griesser *et al.* 2017). Each of the five analyses were performed for 500 simulations of the Brownie algorithm. Sample sizes: (A) n = 127 species, (B) n = 216 species, (C) n = 152 species, (D) n = 156 species, (E) n = 156 species. See Table S6.



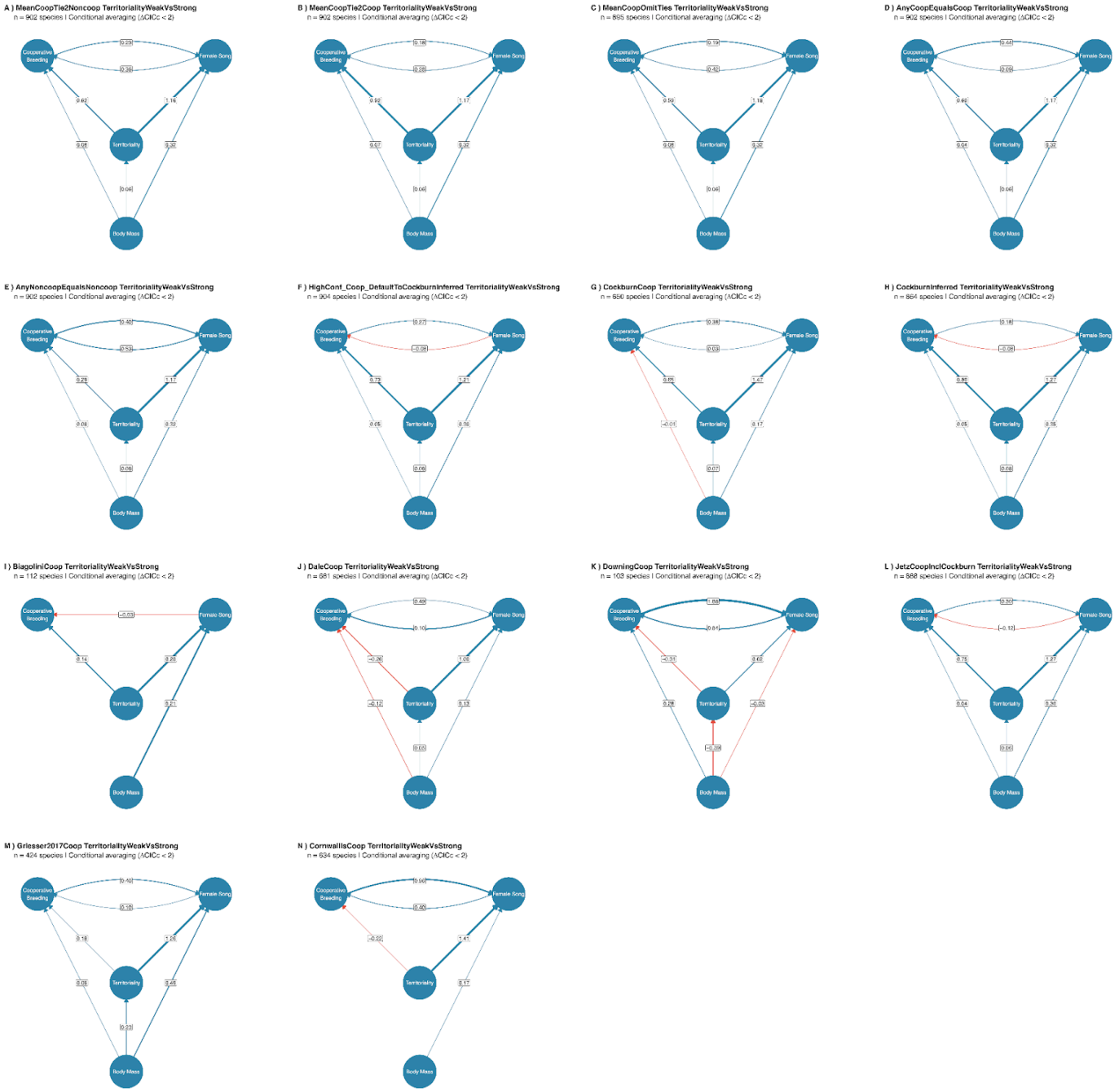
Extended Data Fig. 3: Overlap of stochastic character map simulations of female song and cooperative breeding, measured as the proportion of the tree occupied by each state combination. A) Pooled results of analyses performed on 200 individual trees from BirdTree.org with 20 simulations per tree are shown and correspond with an overall empirical *p*-value of 0.00075. B) Results of analysis performed on the alternate consensus tree for 500 simulations with an empirical *p*-value of 0.008. A red asterisk indicates that at least 95% of the values from the randomized simulations were less than the median value of the simulations of ancestral history based on the true data. (n = 1041 species) See Table S8.


Simulations with variations of the territoriality classification
Transition counts within non-/weakly/seasonally territorial lineages
Transition counts within year-round territorial lineages
A
Non–Cooperative Female Song Absent 23% of tree
Non–Cooperative Female Song Present 30.9% of tree
Cooperative Female Song Absent 1.3% of tree
Cooperative Female Song Present 2.6% of tree
B
Non–Cooperative Female Song Absent 6.2% of tree
Non–Cooperative Female Song Present 29.2% of tree
Cooperative Female Song Absent 0.9% of tree
Cooperative Female Song Present 5.8% of tree
Median Difference from Expected # of Transitions
n.s.
Simulations with variations of the phylogeny
C
Non–Cooperative Female Song Absent 32.7% of tree
Non–Cooperative Female Song Present 56.7% of tree
Cooperative Female Song Absent 2.5% of tree
Cooperative Female Song Present 7.9% of tree
D
Non–Cooperative Female Song Absent 29.5% of tree
Non–Cooperative Female Song Present 60% of tree
Cooperative Female Song Absent 2.5% of tree
Cooperative Female Song Present 8.1% of tree
Simulations with variations of the cooperative breeding classification scheme
E
Non–Cooperative Female Song Absent 28.7% of tree
Non–Cooperative Female Song Present 58.1% of tree
Cooperative Female Song Absent 3% of tree
Cooperative Female Song Present 10% of tree
F
Non–Cooperative Female Song Absent 28.8% of tree
Non–Cooperative Female Song Present 58.6% of tree
Cooperative Female Song Absent 2.9% of tree
Cooperative Female Song Present 9.5% of tree
G
Non–Cooperative Female Song Absent 28.6% of tree
Non–Cooperative Female Song Present 59.1% of tree
Cooperative Female Song Absent 2.7% of tree
Cooperative Female Song Present 9.3% of tree
H
Non–Cooperative Female Song Absent 28.3% of tree
Non–Cooperative Female Song Present 56.8% of tree
Cooperative Female Song Absent 3.4% of tree
Cooperative Female Song Present 11.3% of tree
I
Non–Cooperative Female Song Absent 29.5% of tree
Non–Cooperative Female Song Present 60% of tree
Cooperative Female Song Absent 2.2% of tree
Cooperative Female Song Present 8.2% of tree

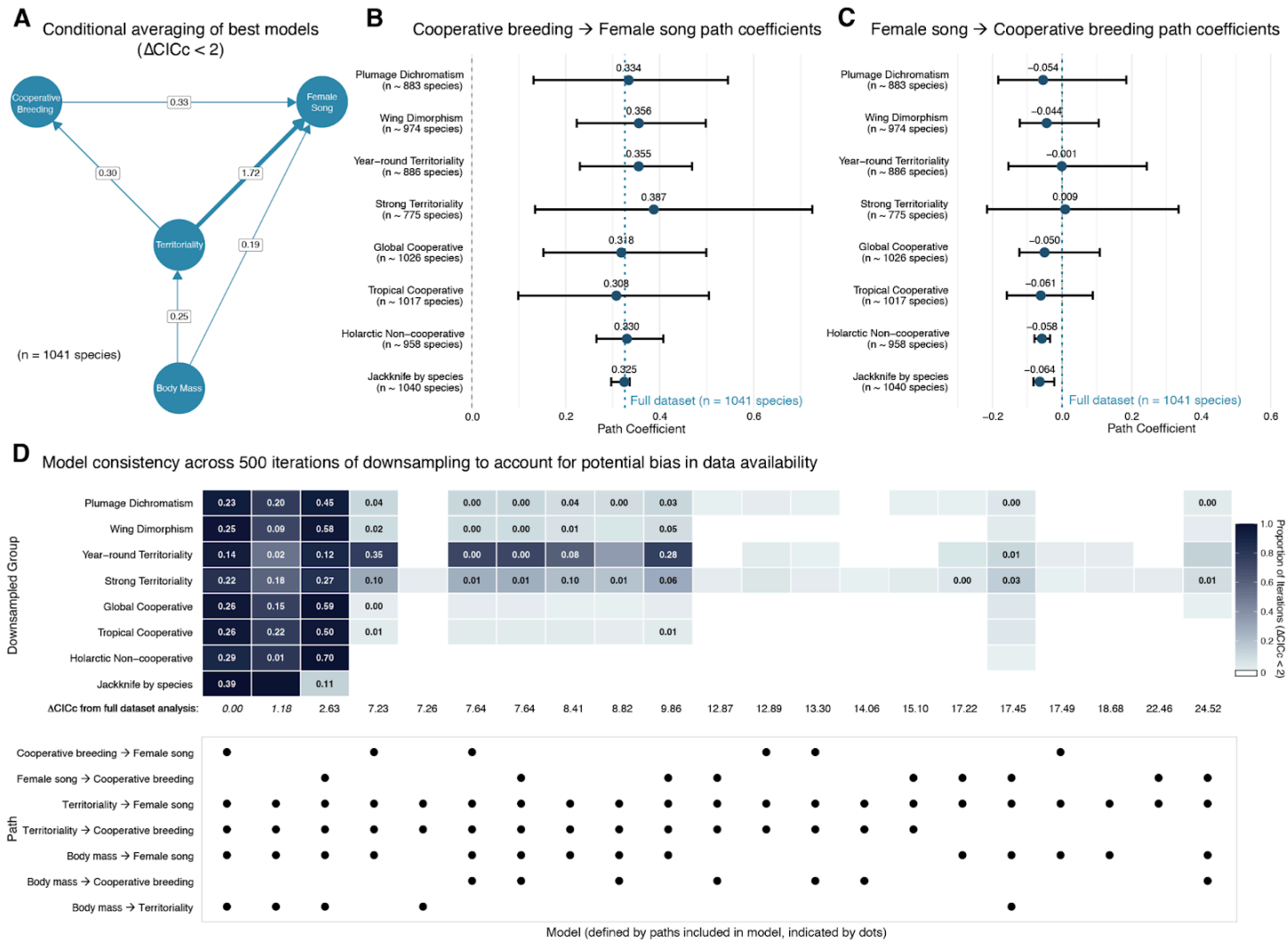
Extended Data Fig. 4: Female song and cooperative breeding transition likelihoods based on the median difference from expected transition counts, calculated using alternate territory classification, alternative phylogenies to account for phylogenetic uncertainty, or alternative cooperative breeding classification methods. Using the default cooperative breeding classification method: Panels A and B show the numbers of transitions between female song states and cooperative breeding states that occurred in lineages that were A) non-territorial or weakly/seasonally territorial and B) year-round territorial. C) Pooled results over 200 trees, 20 simulations per tree. D) Results from 500 simulations using the real trait distribution on the consensus phylogeny obtained from using the “mean.edge” method with absent edges ignored in the mean edge length calculations. Using the default consensus tree: E) MeanCoopTie2Noncoop, F) MeanCoopTie2Coop, G) MeanCoopOmitTies, H) AnyCoopEqualsCoop, and I) AnyNoncoopEqualsNoncoop. Arrow weights indicate the log-transformed median relative number of state switches that occurred on that topography. Gray arrows indicate rates for which the observed number of transitions were not significantly different from expected based on an ANOVA. Otherwise, color denotes the median difference from the expected number of transitions across simulations, with red indicating that the observed number of state switches was greater than expected, and blue indicating fewer state switches than expected. Asterisks indicate that the percent of simulations where the number of transitions is greater (red arrows) or less (blue arrows) than the calculated expected number of transitions for the simulation in question in >95% of simulations (**) or 90-95% of simulations (*). Sample sizes: A) n = 637 species, B) n = 404 species, C) n = 1041 species, D) n = 1041 species, E) n = 1075 species, F) n = 1075 species, G) n = 1066 species, H) n = 1075 species, I) n = 1075 species.



Extended Data Fig. 5: Overlap of stochastic character map simulations of female song and categorical social traits, measured as the proportion of the tree occupied by each state combination: A) familial living (based on data from [16] and female song (empirical *p*-value 0.092, 500 simulations, N = 468 species), and B) social system classification, encompassing both cooperative breeding and familial living[16], and female song (empirical *p*-value 0.0173, 1500 simulations, N = 468 species). Results of post-hoc analyses comparing real and randomized simulations are shown in panel B; a red asterisk indicates that at least 95% of the values from the randomized simulations were less than the median value of the simulations of ancestral history based on the true data, while a blue asterisk indicates that at least 95% of the randomized simulation values were higher than the median value of the simulations of ancestral history based on the true data (see Table ED2, Table S11).



Extended Data Fig. 6: Directed acyclic graphs showing outputs of phylopath analyses with different methods of cooperative breeding classification (A-F) and cooperative breeding classifications based on single sources (G-N).



Extended Data Fig. 7: Robustness of the cooperative breeding-female song association to data availability biases using phylogenetic path analysis, with year-round territoriality and body mass as additional factors. (A) Directed acyclic graph showing conditional-averaged path coefficients from phylogenetic path analysis of the relationships between female song presence, cooperative breeding, year-round territoriality, and body mass (log-scaled and normalized). This figure can be compared to the Fig. 4 main text, which shows phylogenetic path analyses with strength of territoriality included instead of year-round territoriality shown here. Arrow thickness indicates the magnitude of path coefficients, with all paths shown representing positive relationships. Only models within ΔCICc < 2 (2 models) were included in the averaging. (B) Forest plot showing the mean and 95% confidence interval for the cooperative breeding → female song path coefficient across 500 iterations for each bias correction type, compared to the value from the full dataset. (C) Forest plot showing the mean and 95% confidence interval for the female song → cooperative breeding path coefficient across 500 iterations for each bias correction type, compared to the value from the full dataset. (D) Model consistency across bias corrections. Heatmap showing the number of iterations for which each phylogenetic path model was significant (had ΔCICc < 2) across 500 iterations for each bias correction type. The row underneath the heatmap contains the ΔCICc values for each model in the phylopath analysis performed using the full database, with the ΔCICc values of the two models that were statistically equivalent (ΔCICc < 2) in the full dataset in italics. The lower panel shows the models based on the presence or absence of each rate in the model. The models shown are all those that were in the top 20 models by # of iterations with ΔCICc < 2 for any of the bias correction types.

Social Trait (Trait B)	Number of species	Empirical p-value	Degree of association between Trait A (female song absent versus female song present) and Trait B			
			Female song absent co-occurrence with Social Trait = 0	Female song present co-occurrence with Social Trait = 0	Female song absent co-occurrence with Social Trait = 1	Female song present co-occurrence with Social Trait = 1
Asocial (0) vs. social groups (1)	255	0.438	0.432	0.36	0.548	0.608
Asocial or pair (0) vs larger group (1)	255	0.288	0.658	0.454	0.442	0.528
Asocial, pair, or small group (0) vs. large group (1)	255	0.316	0.392	0.464	0.772	0.494
Social bonds lasting one season or less (0) vs longer than one breeding season (1)	270	0.126	0.522	0.012	0.552	1
Non-familial (0) vs. familial living (1)	468	0.092	0.82	0.028	0.568	0.908
Non-colonial (0) vs. colonial living (1)	270	0.374	0.596	0.294	0.552	0.684
Monogamy (0) vs. polygyny (1)	323	0.248	0.128	0.716	0.758	0.62
One or two caretakers (0) vs. more than two caretakers (1)	270	0.156	0.586	0.214	0.398	0.968
Weak or no territoriality (0) vs. strong territoriality	919	0.002	0.992	0.010	0.260	0.706
Weak, seasonal, or no territoriality (0) vs. year-round territoriality (1)	1091	<0.002	0.998	0.056	0	0.986
Non-cooperative (0) vs. cooperative breeding (1)	1041	0.002	0.274	0.194	0.504	1

Extended Data Table 1: Results of overlapping simulations of ancestral character history for binary social variables and female song. For each pair of traits (female song + a social variable from the left column), we ran 500 ancestral character mapping simulations with the real data and 500 with shuffled data (i.e. with tip states in a randomized order) for comparison. We used methods adapted from[28] to estimate an empirical p-value (yellow indicates significance at $p < 0.05$). Across the Trait A-Trait B stochastic character map pairs based on the real data, we found the median proportion of the phylogeny that each combination of traits occupied, then approximated the degree of association between trait states by comparing the proportion of the tree occupying a given trait combination in the the real versus randomized simulations. Cells highlighted in blue indicate that that trait combination was less than the median of real simulations in <5% of randomized simulations (and thus the trait combination is less common than expected if traits were independent). Cells highlighted in red indicate that >95% of randomized simulations were less than the median of the real simulations (thus the trait combination is more common than expected if traits were independent).

References

1. Catchpole, C. K. & Slater, P. J. B. *Bird Song: Biological Themes and Variations*. (Cambridge University Press, 2003).
2. Odom, K. J., Hall, M. L., Riebel, K., Omland, K. E. & Langmore, N. E. Female song is widespread and ancestral in songbirds. *Nat. Commun.* **5**, 3379 (2014).
3. Riebel, K., Odom, K. J., Langmore, N. E. & Hall, M. L. New insights from female bird song: towards an integrated approach to studying male and female communication roles. *Biol. Lett.* **15**, 20190059 (2019).
4. Langmore, N. E. Functions of duet and solo songs of female birds. *Trends Ecol. Evol.* **13**, 136–140 (1998).
5. Hall, M. L. A review of hypotheses for the functions of avian duetting. *Behav. Ecol. Sociobiol.* **55**, 415–430 (2004).
6. Barros, C. M., Benedict, L. & Sanchez, K. A. A review of the literature on female birdsong function. *Anim. Behav.* **216**, 23–35 (2024).
7. Tobias, J. A., Montgomerie, R. & Lyon, B. E. The evolution of female ornaments and weaponry: social selection, sexual selection and ecological competition. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **367**, 2274–2293 (2012).
8. West-Eberhard, M. J. Sexual selection, social competition, and speciation. *Q. Rev. Biol.* **58**, 155–183 (1983).
9. Earl, A. D. *et al.* A cryptic role for reciprocal helping in a cooperatively breeding bird. *Nature* **642**, 381–388 (2025).
10. Cockburn, A. Prevalence of different modes of parental care in birds. *Proc. Biol. Sci.* **273**, 1375–1383 (2006).
11. Rubenstein, D. R. & Lovette, I. J. Reproductive skew and selection on female ornamentation in social species. *Nature* **462**, 786–789 (2009).
12. Clutton-Brock, T. H. *et al.* Intrasexual competition and sexual selection in cooperative mammals. *Nature* **444**, 1065–1068 (2006).
13. Ben Mocha, Y., Dahan, T., Zou, Y., Griesser, M. & Markman, S. Evidence for a reproductive sharing

continuum in cooperatively breeding mammals and birds: consequences for comparative research. *Proc. Biol. Sci.* **290**, 20230607 (2023).

14. Apakupakul, K. & Rubenstein, D. R. Bateman's principle is reversed in a cooperatively breeding bird. *Biol. Lett.* **11**, 20150034 (2015).
15. Hauber, M. E. & Lacey, E. A. Bateman's Principle in Cooperatively Breeding Vertebrates: The Effects of Non-breeding Alloparents on Variability in Female and Male Reproductive Success. *Integr. Comp. Biol.* **45**, 903–914 (2005).
16. Griesser, M., Drobniak, S. M., Nakagawa, S. & Botero, C. A. Family living sets the stage for cooperative breeding and ecological resilience in birds. *PLoS Biol.* **15**, e2000483 (2017).
17. Dickinson, J. L. & Hatchwell, B. J. Fitness consequences of helping. *Ecology and evolution of cooperative breeding in birds* 48–66 (2004).
18. Hatchwell, B. J. The evolution of cooperative breeding in birds: kinship, dispersal and life history. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **364**, 3217–3227 (2009).
19. Ligon, J. D. & Burt, D. B. Evolutionary origins. in *Cooperative breeding in birds* (eds. Koenig, W. D. & Dickinson, J. L.) 5–34 (Cambridge University Press, Cambridge, UK, 2004).
20. Shen, S.-F., Emlen, S. T., Koenig, W. D. & Rubenstein, D. R. The ecology of cooperative breeding behaviour. *Ecol. Lett.* **20**, 708–720 (2017).
21. Tobias, J. A. *et al.* Territoriality, social bonds, and the evolution of communal signaling in birds. *Front. Ecol. Evol.* **4**, (2016).
22. Leighton, G. M. Cooperative breeding influences the number and type of vocalizations in avian lineages. *Proc. Biol. Sci.* **284**, (2017).
23. Odom, K. J. *et al.* Sex role similarity and sexual selection predict male and female song elaboration and dimorphism in fairy-wrens. *Ecol. Evol.* **11**, 17901–17919 (2021).
24. Odom, K. J. *et al.* Global incidence of female birdsong is predicted by territoriality and biparental care in songbirds. *Nat. Commun.* **16**, 6157 (2025).
25. Griesser, M., Drobniak, S. M., Graber, S. M. & van Schaik, C. P. Parental provisioning drives brain size in birds. *Proc. Natl. Acad. Sci. U. S. A.* **120**, e2121467120 (2023).

26. O'Meara, B. C., Ané, C., Sanderson, M. J. & Wainwright, P. C. Testing for different rates of continuous trait evolution using likelihood. *Evolution* **60**, 922–933 (2006).

27. Jetz, W., Thomas, G. H., Joy, J. B., Hartmann, K. & Mooers, A. O. The global diversity of birds in space and time. *Nature* vol. 491 444–448 Preprint at https://doi.org/10.1038/nature11631 (2012).

28. Huelsenbeck, J. P., Nielsen, R. & Bollback, J. P. Stochastic mapping of morphological characters. *Syst. Biol.* **52**, 131–158 (2003).

29. Clutton-Brock, T. H. Cooperative breeding in mammals. in *Cooperation in Primates and Humans: Mechanisms and Evolution* (eds. Kappeler, P. M. & van Schaik, C. P.) 173–190 (Springer Berlin Heidelberg, Berlin, Heidelberg, 2006).

30. Yamamoto, M. E. *et al.* Male and female breeding strategies in a cooperative primate. *Behav. Processes* **109 Pt A**, 27–33 (2014).

31. Eens, M. & Pinxten, R. Female European starlings increase their copulation solicitation rate when faced with the risk of polygyny. *Anim. Behav.* **51**, 1141–1147 (1996).

32. Sandell, M. I. & Smith, H. G. Female aggression in the European starling during the breeding season. *Anim. Behav.* **53**, 13–23 (1997).

33. Snyder, K. T. & Creanza, N. Polygyny is linked to accelerated birdsong evolution but not to larger song repertoires. *Nat. Commun.* **10**, 884 (2019).

34. Riehl, C. Evolutionary routes to non-kin cooperative breeding in birds. *Proc. Biol. Sci.* **280**, 20132245 (2013).

35. Rubenstein, D. R. Family feuds: social competition and sexual conflict in complex societies. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **367**, 2304–2313 (2012).

36. Riters, L. V., Polzin, B. J., Maksimoski, A. N., Stevenson, S. A. & Alger, S. J. Birdsong and the Neural Regulation of Positive Emotion. *Front. Psychol.* **13**, 903857 (2022).

37. Maksimoski, A. N. *et al.* The motivation to flock correlates with vocal-social behaviors and dopamine-related gene expression in male European starlings. *Horm. Behav.* **153**, 105374 (2023).

38. Loning, H., Griffith, S. C. & Naguib, M. The ecology of zebra finch song and its implications for vocal communication in multi-level societies. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **379**, 20230191 (2024).

39. Odom, K. J. *et al.* Typical Males and Unconventional Females: Songs and Singing Behaviors of a Tropical, Duetting Oriole in the Breeding and Non-Breeding Season. *Frontiers in Ecology and Evolution* **4**, (2016).

40. Halkin, S. L. Nest-vicinity song exchanges may coordinate biparental care of northern cardinals. *Anim. Behav.* **54**, 189–198 (1997).

41. Ritchison, G. The Function of Singing in Female Black-Headed Grosbeaks (Pheucticus melanocephalus): Family-Group Maintenance. *Auk* **100**, 105–116 (1983).

42. Payne, R. B., Payne, L. L., Rowley, I. & Russell, E. M. Social Recognition and Response to Song in Cooperative Red-Winged Fairy-Wrens. *Auk* **108**, 811–819 (1991).

43. Payne, R. B., Payne, L. L. & Rowley, I. Kin and social relationships in splendid fairy-wrens: recognition by song in a cooperative bird. *Anim. Behav.* **36**, 1341–1351 (1988).

44. Haven Wiley, R. & Wiley, M. S. Recognition of Neighbors' Duets By Stripe-Backed Wrens Campylorhynchus Nuchalis. *Behaviour* **62**, 10–34 (1977).

45. Hausberger, M. & Black, J. M. Female song in European starlings: the case of non-competitive song-matching. *Ethol. Ecol. Evol.* **3**, 337–344 (1991).

46. Leedale, A. E. *et al.* Cost, risk, and avoidance of inbreeding in a cooperatively breeding bird. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 15724–15730 (2020).

47. Snyder, K. T., Sellers, M. L. & Creanza, N. Cultural shifts after punctuated environmental stress: a study of song distributions in Dark-eyed Junco and Song Sparrow populations. *J. Field Ornithol.* (2024) doi:10.5751/jfo-00442-950203.

48. Rubenstein, D. R. Spatiotemporal environmental variation, risk aversion, and the evolution of cooperative breeding as a bet-hedging strategy. *Proc. Natl. Acad. Sci. U. S. A.* **108 Suppl 2**, 10816–10822 (2011).

49. Camerlenghi, E., Nolazco, S., Farine, D. R., Magrath, R. D. & Peters, A. Social restructuring during harsh environmental conditions promotes cooperative behaviour in a songbird. *Proc. Biol. Sci.* **291**, 20232427 (2024).

50. Jetz, W. & Rubenstein, D. R. Environmental uncertainty and the global biogeography of cooperative breeding in birds. *Curr. Biol.* **21**, 72–78 (2011).

51. Rubenstein, D. R. & Lovette, I. J. Temporal environmental variability drives the evolution of cooperative

957 breeding in birds. *Curr. Biol.* **17**, 1414–1419 (2007).

958 52. Cornwallis, C. K. *et al.* Cooperation facilitates the colonization of harsh environments. *Nat Ecol Evol* **1**, 57
959 (2017).

960 53. Downing, P. A., Cornwallis, C. K. & Griffin, A. S. Sex, long life and the evolutionary transition to
961 cooperative breeding in birds. *Proc. Biol. Sci.* **282**, 20151663 (2015).

962 54. Langen, T. A. Prolonged offspring dependence and cooperative breeding in birds. *Behav. Ecol.* **11**,
963 367–377 (2000).

964 55. MacDougall-Shackleton, S. A. Developmental stress and birdsong: integrating signal function and
965 development. *Current Opinion in Behavioral Sciences* **6**, 104–110 (2015).

966 56. Griesser, M. & Suzuki, T. N. Occasional cooperative breeding in birds and the robustness of comparative
967 analyses concerning the evolution of cooperative breeding. *Zoological Lett* **2**, 7 (2016).

968 57. Cockburn, A. Chapter Six - Can't see the 'hood' for the trees: Can avian cooperative breeding currently be
969 understood using the phylogenetic comparative method? in *Advances in the Study of Behavior* (eds.
970 Naguib, M. et al.) vol. 52 243–291 (Academic Press, 2020).

971 58. Dunn, P. O., Armenta, J. K. & Whittingham, L. A. Natural and sexual selection act on different axes of
972 variation in avian plumage color. *Sci Adv* **1**, e1400155 (2015).

973 59. Biagolini, C., Westneat, D. F. & Francisco, M. R. Does habitat structural complexity influence the frequency
974 of extra-pair paternity in birds? *Behav. Ecol. Sociobiol.* **71**, (2017).

975 60. Dale, J., Dey, C. J., Delhey, K., Kempenaers, B. & Valcu, M. The effects of life history and sexual selection
976 on male and female plumage colouration. *Nature* **527**, 367–370 (2015).

977 61. Tobias, J. A. *et al.* AVONET: morphological, ecological and geographical data for all birds. *Ecol. Lett.* **25**,
978 581–597 (2022).

979 62. Billerman, S. M., Keeney, B. K., Rodewald, P. G., Schulenberg, T. S. & Others. Birds of the World.
980 https://birdsoftheworld.org/bow/home (2025).

981 63. Webb, W. H. *et al.* Female Song Occurs in Songbirds with More Elaborate Female Coloration and
982 Reduced Sexual Dichromatism. *Frontiers in Ecology and Evolution* **4**, (2016).

983 64. Garamszegi, L. Z., Pavlova, D. Z., Eens, M. & Møller, A. P. The evolution of song in female birds in

Europe. *Behav. Ecol.* **18**, 86–96 (2006).

65. Price, J. J. Evolution and life-history correlates of female song in the New World blackbirds. *Behav. Ecol.* **20**, 967–977 (2009).

66. Snow, D. W. & Cramp, S. *The Complete Birds of the Western Palearctic*. (Oxford University Press, 1998).

67. Mikula, P. *et al.* Female solo song and duetting are associated with different territoriality in songbirds. *Behav. Ecol.* **31**, 322–329 (2019).

68. VertLife : Data downloads (File: HackettStage2_3001_4000.zip). https://data.vertlife.org/?basetree=birdtree&start_folder=Stage2/.

69. Revell, L. J. phytools: an R package for phylogenetic comparative biology (and other things): phytools: R package. *Methods Ecol. Evol.* **3**, 217–223 (2012).

70. Maddison, W. P. & FitzJohn, R. G. The unsolved challenge to phylogenetic correlation tests for categorical characters. *Syst Biol* **64**, 127–136 (2015).

71. Ho, L. S. T. & Ané, C. Intrinsic inference difficulties for trait evolution with Ornstein-Uhlenbeck models. *Methods Ecol. Evol.* **5**, 1133–1146 (2014).

72. Ho, L. si T. & Ané, C. A Linear-Time Algorithm for Gaussian and Non-Gaussian Trait Evolution Models. *Syst Biol* **63**, 397–408 (2014).

73. van der Bijl, W. phylopath: Easy phylogenetic path analysis in R. *PeerJ* **6**, e4718 (2018).