Sexual selection, body mass and molecular evolution interact to predict diversification in birds

Maider Iglesias-Carrasco¹, Michael D. Jennions¹, Simon Y. W. Ho² and David A. Duche ˆne¹,²

¹Research School of Biology, Australian National University, Canberra, Australian Capital Territory 2601, Australia
²School of Life and Environmental Sciences, University of Sydney, Sydney, New South Wales 2006, Australia

Electronic supplementary material is available online at https://dx.doi.org/10.6084/m9.figshare.c.4421279.

1. Introduction

Sexual selection is often invoked as a major driver of speciation or extinction [1], and several comparative analyses have concluded that it elevates the net rate of diversification (speciation minus extinction) [2–4] (but see [5–7]). However, there is limited evidence about the mechanisms underlying this association, partly because the relationship between sexual selection and diversification is often explored in isolation from other well-known evolutionary consequences of sexual selection. For example, the reported association between sexual selection and diversification could have arisen because other traits that evolve in response to sexual selection directly affect rates of speciation and/or extinction. Two such traits that have been repeatedly linked to speciation and extinction rates are the rate of molecular evolution and adult body size [8–10] (figure 1).

There are several ways in which sexual selection could directly influence speciation and extinction. First, female mate choice can drive the evolution of ecologically relevant traits that are genetically correlated with preferred male traits [11,12], increasing positive selection in those parts of the genome that facilitate local adaptation and lead to ecological divergence between populations [11,13–16]. Second, reduced gene flow between populations can arise from the rapid coevolution of male–female sexual traits along different arbitrary trajectories, caused by either mate choice (e.g. Fisherian runaway [11])
Figure 1. Hypothesized causal relationships between sexual selection, the rate of molecular evolution and the net rate of diversification. The red arrow shows the proposed hypothesis that sexual selection acts through other species traits, including those associated with body mass and the rate of molecular evolution, to drive higher rates of speciation or lower rates of extinction.

Figure 1

- **diversification rate**
- **rate of molecular evolution**
- **sexual selection**
- **body mass**

or sexual conflict (‘arms race’ between sexually antagonistic traits [15,17,18]). Third, sexual selection can elevate the risk of extinction by reducing the environmental ‘fit’ of individuals. This might be due to sexual conflict, a trade-off in trait expression between natural and sexual selection [19,20] or a reduction in the effective population size because only some individuals of the competitive sex reproduce [21].

Sexual selection acts on mutations that increase reproductive competitiveness, seductiveness, choosiness and mating resistance. This should increase the rate of evolution in regions of the genome encoding for these functional traits. Furthermore, sexual selection can accelerate the purging of deleterious mutations that affect resource acquisition and assimilation (i.e. condition). This is because these mutations are likely to have detrimental effects on the multitude of ecologically relevant traits that affect the expression of condition-dependent, sexually selected traits [22,23]. This purging reduces genomic changes that are harmful to population persistence, accelerates adaptive evolution [24], increases population fitness [25] and can even lower the costs of sex [23] (but see [26]). Sexual selection might also indirectly affect the type and number of mutations that become fixed if it lowers the effective population size, which leads to stronger genetic drift and a reduced efficacy of selection for beneficial mutations (but see [22]).

Another mechanism by which sexual selection can influence diversification is to favour a higher mutation rate. Directional female mate choice for exaggerated male traits is expected to reduce additive genetic variation, thereby decreasing the benefits of choice. However, this reduced genetic diversity is not observed in nature (the ‘lek paradox’ [27]). One explanation for this is an increased input of novel genetic variation through a higher mutation rate [28]. Theory suggests that ‘mutator’ alleles can spread if nonlinear effects of increased trait expression (i.e. larger traits disproportionately increase fitness) select for greater phenotypic benefits of increased trait expression (i.e. larger traits disproportional increases in testes and is prevalent in most species under strong sexual selection, leads to greater accumulation of copy errors in the germline [9]. Sexual selection could be linked to diversification rates if an elevated mutation rate accelerates the appearance of genetic incompatibilities between populations [29].

Sexual selection therefore has several opposing effects on genome evolution that might each generate predictable, but opposing, macroevolutionary patterns [22]. In the light of the processes already mentioned, one possible long-term effect of sexual selection is an accelerated mutation rate, which facilitates adaptation because of greater production of beneficial mutations. Species with a history of stronger sexual selection might then be more likely to persist due to faster molecular evolution in regions of the genome that are under natural selection. Alternatively, these regions might show reduced molecular evolution if the predominant effect of sexual selection is not to elevate the mutation rate but simply to eliminate slightly deleterious mutations. Or it might be the case that sexual selection has a negligible effect on fixation of either beneficial or detrimental mutations and that solely by elevating the mutation rate [28] it drives genetic incompatibilities associated with promoting speciation.

There has been a limited success in disentangling how molecular evolution, sexual selection and other life-history traits affect species diversification. Although numerous traits have been proposed as drivers of both species diversification and molecular evolution, body size can be singled out as a key variable that affects most life-history traits and demographic parameters, including generation time, lifespan and population size. These, in turn, are all factors that are believed to affect rates of molecular evolution [30]. Indeed, there is evidence that smaller body size is positively associated with species richness in many taxa [31]. To date, however, there have been few attempts to determine if, and how, sexual selection, rates of molecular evolution and key life-history traits interact to affect species diversification.

Here, we test whether body size, sexual selection and molecular evolution interact to predict variation in rates of species diversification among bird genera. We collated data on diversification rates, body mass and sexual selection using a measure of the degree of polygyny as a proxy for sexual selection. We estimated non-synonymous (dNS) and synonymous (dS) rates of molecular evolution, and their ratio (dNS/dS), using both nuclear and mitochondrial sequence data (n = 173 and 354 genera, respectively). Then, we tested whether the degree of polygyny (i.e. sexual selection) is associated with net diversification rates while taking into account any interactive effects of the rate of molecular evolution or body mass. We predicted that avian genera with a high degree of polygyny, small body size and higher rates of molecular evolution would have the highest rates of diversification. We tested for the effects of molecular rates that are associated with selection (dNS and dS) or the mutation rate (dNS and dS), or both.

2. Material and methods

(a) Data collection

We compiled information about the degree of polygyny and body mass in 954 species of birds in 558 genera, using existing
data compendia [32,33] (electronic supplementary material). We used data about the degree of polygyny (i.e. variance in mating success of males) in which species were ranked from 1 to 5 [32,34]; 1 = polyandry; 2 = monogamy (less than 5% polygyny); 3 = mostly monogamy, but occasional polygyny (5–15% polygyny); 4 = mostly polygyny (greater than 15% polygyny); and 5 = lekking or promiscuous. This ordinal variable indicates the potential for sexual selection due to male–male competition for mates. For the statistical analyses, however, we used the polygyny score as a continuous variable which was the average value for each avian genus examined (see below). We used an estimate of the complete avian phylogeny and estimated the net diversification rate of every bird species using the reciprocal of the equal splits measure of diversity (strongly correlated with other metrics of net diversification rates; see electronic supplementary material, figure S1) [35]. We then identified all of the avian genera for which we had data for polygyny, body mass and nucleotide sequences of four mitochondrial (n = 354) and four nuclear genes (n = 173; GenBank accessions sourced from Jetz et al. [36]). The data on the degree of polygyny and body mass were averaged across the species within each genus.

To estimate rates of molecular evolution, we aligned the nucleotide sequences across genera (electronic supplementary material, table S9) and estimated \(d_S\) and \(d_N\) substitutions per lineage using the software HyPhy v2.2 [37]. We used the MC94 codon evolution model [38], paired with a general time-reversible model of nucleotide substitution selected automatically using the Akaike information criterion. The terminal branch lengths inferred in this procedure represent genus-average estimates of \(d_S\) and \(d_N\). In order to avoid making misleading estimates of relative molecular evolutionary rates, we removed any genera for which terminal branch length was zero (electronic supplementary material). Estimates of \(d_S\) and \(d_N\) on terminal branches are confounded by the elapsed time over which taxa have been able to accumulate nucleotide changes [39], so we estimated relative rates of \(d_N\) and \(d_S\) evolution for each genus using a model that relaxes the molecular-clock assumption using penalized likelihood [40] (electronic supplementary material). Throughout the text, we differentiate molecular rates that are associated with selection (\(d_N\) and \(\omega\)) and with the mutation rate (\(d_S\) and \(d_0\)). To clarify, \(d_S\) estimates the rate of silent substitutions that do not affect amino acid coding. It is mainly influenced by the mutation rate. By contrast, \(d_N\) is an estimation of nucleotide changes that affect amino acid coding and is influenced by both the mutation rate and selection. In consequence, \(\omega\) is expected to be influenced by selection but not by the mutation rate [8].

### (b) Statistical analyses

We used phylogenetic generalized least squares (PGLS) to examine the relationships between the degree of polygyny (calculated as the mean of the available species values in each genus), diversification rate, relative rate of molecular evolution and body mass (see electronic supplementary material for information on the calculation of scores). We used the avian tree containing the sampled genera to correct for phylogenetic non-independence. In addition, we investigated how phylogenetic uncertainty affected our results. Phylogenetic regression models were replicated with each of 100 randomly selected trees from a published posterior distribution of avian relationships [36]. To linearize relationships, we log-transformed the rates of diversification, rates of molecular evolution and body mass (all continuous variables).

In PGLS, we assumed a lambda model of trait evolution [41], which provided a better fit than three other commonly used models (Brownian motion, Kappa and an Ornstein–Uhlenbeck process). The parameters of the PGLS regression models were optimized using maximum likelihood in the nlme R package [42], and model residuals were inspected for normality.

Regression models that included two-way and three-way interactions were re-run with interactions excluded. If removal did not reduce the model fit (as determined by a likelihood-ratio test), we interpreted parameter estimates from the reduced model. Otherwise, we interpreted the estimates of retained interaction terms (electronic supplementary material, tables S1–S5).

We ran PGLS separately for mitochondrial and nuclear genes to address the two main questions described below. The links between these questions are illustrated in figure 1. We then explored the directionality of these relationships using phylogenetic path analysis [43]. We based our four hypotheses of path models (electronic supplementary material, figure S2) on a previous study of the association between life history, molecular evolution rate and species richness in angiosperms [44].

(i) Do the degree of polygyny and body mass predict the rate of molecular evolution?

We treated the degree of polygyny, body mass and their two-way interaction as fixed factors, and tested whether they explain variation in \(d_S\), \(d_N\) or \(\omega\). When testing for the influence of the degree of polygyny and body mass on \(d_S\), we also repeated the model with \(d_S\) as a covariate. This equates to exploring the effect of the degree of polygyny and body mass on \(d_N\) independent of the influence of the mutation rate (estimated as \(d_0\)).

(ii) Do the rate of molecular evolution, degree of polygyny and body size interact to predict the rate of diversification?

Molecular evolution, the degree of polygyny and body mass are likely to be related to each other and might have interactive effects on net rates of diversification (figure 1). We therefore ran models that tested whether the molecular rates \(d_S\), \(d_N\) or \(\omega\), in conjunction with the degree of polygyny, body mass and all two-way and three-way interactions, explain variation in the net rate of diversification.

### 3. Results

(a) Do the degree of polygyny and body mass predict the rate of molecular evolution?

A higher rate of synonymous mutation \(d_S\) for mitochondrial genes was associated with a higher mean score for polygyny (estimate ± s.e. = 0.097 ± 0.036, \(p = 0.008\)), but not with body mass or the interaction between body mass and the degree of polygyny (electronic supplementary material, table S1). For nuclear genes, there was no effect of body mass, degree of polygyny or their interaction on \(d_S\).

For the rate of non-synonymous mutation \(d_N\) for mitochondrial genes, the PGLS model showed no effect of body mass, degree of polygyny or their interaction (electronic supplementary material, table S2). By contrast, \(d_N\) for nuclear genes was positively associated with the degree of polygyny (estimate ± s.e. = 0.215 ± 0.099, \(p = 0.031\); electronic supplementary material, table S2), but not with body mass nor with an interaction between the degree of polygyny and body mass. When we then ran models that included \(d_N\) as a covariate, it was the main predictor of \(d_N\) (mitochondrial: estimate ± s.e. = 0.442 ± 0.041, \(p < 0.001\); nuclear: estimate ± s.e. = 0.685 ± 0.094, \(p < 0.001\); electronic supplementary material, figure S3); the degree of polygyny, body mass and their interaction failed to explain any additional variation. The use of path analysis
supported our conclusion that $d_S$ is the main driver of $d_N$ in nuclear genes (figure 2), but not in mitochondrial genes. There was no evidence that $\omega$ for either mitochondrial or nuclear genes is associated with the degree of polygyny, body mass or their interaction (electronic supplementary material, table S4).

(b) Do the rate of molecular evolution, degree of polygyny and body mass interact to predict the rate of diversification?

The net rate of diversification was explained by a three-way interaction between the degree of polygyny, body mass and $d_N$ for both mitochondrial and nuclear genes (mitochondrial interaction estimate $\pm$ s.e. $= 0.044 \pm 0.022$, $p = 0.039$; nuclear interaction estimate $\pm$ s.e. $= 0.05 \pm 0.022$, $p = 0.024$). The same interaction was present for nuclear $d_S$ (interaction estimate $\pm$ s.e. $= 0.129 \pm 0.043$, $p = 0.003$; figure 3), but was absent for mitochondrial $d_{N_0}$ and for either mitochondrial or nuclear estimates of $\omega$. The significant three-way interactions for mitochondrial and nuclear $d_N$ and nuclear $d_S$ all showed a positive association between the degree of polygyny and the diversification rate in small-bodied taxa that was unaffected by the rate of molecular evolution (electronic supplementary material, table S5). In large-bodied taxa, there was a positive association between the degree of polygyny and diversification rates, but only when the rate of molecular evolution was high.

In models where the three-way interaction was not significant and was removed, the diversification rate was associated with an interaction between the degree of polygyny and body mass (electronic supplementary material, table S5). The single exception occurred for the nuclear $\omega$,
which was positively associated with diversification rate (estimate ± s.e. = 0.1 ± 0.04, \( p = 0.014 \)). We found that the results of replicate phylogenetic regression analyses on a sample of 100 trees were consistent with those using the published summary tree estimate (electronic supplementary material, figure S4). One exception was the analyses of the three-way interaction for mitochondrial \( d_s \), which led to \( p \)-values above 0.05 in 24% of posterior trees.

Phylogenetic path analyses supported our finding of a three-way interaction between the degree of polygyny, body mass and \( d_s \) (but not \( d_S \)) for nuclear genes (figure 2), but not for mitochondrial genes. In sum, our data are broadly consistent with the hypothesis that the effect of the degree of polygyny on diversification is mediated by mutation rates (\( d_N \) and \( d_S \)) and body mass (or life-history traits that scale with body mass).

4. Discussion

Our data provide evidence that high degrees of polygyny are associated with an increased rate of net diversification across bird genera with small average body size. This finding of an association between sexual selection and net diversification is consistent with theoretical expectations [15], and with the results of previous studies (e.g. [2–4]). More interestingly, we found that molecular evolutionary rates influenced by the mutation rate (\( d_N \) and \( d_S \)) interact with the degree of polygyny and body mass to explain some of the variation in diversification rates in large-bodied taxa. The degree of polygyny we used for our analysis is related to the social mating system of species and therefore represents a good proxy for the strength of sexual selection [34]. However, it should be noted that the degree of polygyny is correlated with other species traits, such as the degree of parental care (see electronic supplementary material).

The primary finding of our study is that the association between net diversification rates and the degree of polygyny is mediated by body mass and the rate of molecular evolution. We found that the interaction between the degree of polygyny and body mass explains net diversification rates in phylogenetic regression models (electronic supplementary material, table S5). However, an association between polygyny and net diversification rate was only general in small-bodied genera. Among large-bodied genera, it is only those with high rates of molecular evolution that show a positive association between the degree of polygyny and net diversification rate (electronic supplementary material, table S5). One explanation for this pattern is that the benefits of sexual selection in facilitating local adaptation are greater in smaller-bodied genera. Smaller species tend to have larger effective population sizes, which are expected to increase the likelihood that beneficial mutations reach fixation [45]. Alternatively, larger-bodied taxa might be more reliant than smaller-bodied taxa on mechanisms that increase their genetic variability: larger-bodied organisms generally have greater long-term susceptibility to extinction in the face of environmental change [46,47]. This greater susceptibility might be partly due to their lower rate of molecular evolution (per unit time). Slower speciation or increased extinction in some larger-bodied taxa might therefore be caused by the combination of smaller population size and lower reproductive potential [48], and reduced genetic diversity due to sexual selection [27] unless compensated for by a higher mutation rate [28].

Strikingly, we found that the association between net diversification rates and the interaction between the degree of polygyny, body mass and molecular evolution was mediated by the mutation rate (estimated as \( d_s \) and \( d_S \)) rather than by the strength of selection (estimated as \( d_N \) and \( w_o \)). One explanation for the lack of a mediating effect of molecular evolutionary rates attributed to selection is that their impact might only be detectable in specific genes, such as those coding for sexual traits [49]. Future explorations of molecular evolutionary rates in those regions of the genome will provide additional insights into the relative impact of sexual selection across different genomic regions (e.g. faster evolution of seminal-protein genes [50]). Nonetheless, our results are broadly consistent with our hypothesis that an elevated mutation rate in taxa with strong sexual selection can accelerate reproductive isolation. Despite natural selection across metazoans to reduce the mutation rate [51], sexual selection due to female mate choice has controversially been suggested to adaptively increase the mutation rate to elevate phenotypic variability in traits with nonlinear fitness returns [9,28]. This might manifest as sexual selection enhancing the spread of mutator alleles that hitch-hike with beneficial mutations [28] that increase the probability of adaptation and the resilience of populations to environmental change. It should be noted, however, that the degree of polygyny was only significantly associated with the estimated mutation rate (\( d_S \)) in mitochondrial genes, and not in nuclear genes (question (i); electronic supplementary material, table S1). Consequently, we cannot reject the hypothesis that the association seen in mitochondrial genes is driven by another process. Finally, we found that the direction of the observed associations in several of our models was similar for both nuclear and mitochondrial genes, lending some robustness to our findings. Sexual selection is expected to have a stronger effect in nuclear genes coding for traits involved in male–male competition and mate choice. Similarly, mitochondrial genomes are subject to a wide range of selective constraints that might influence estimates of \( d_N \) and \( d_S \). One noteworthy example is the preference for codons ending in A and T in many protein-coding regions [52], which can decrease the power to estimate proxies of mutation rate accurately. This is also the case with other influences on molecular rate estimates that might overshadow the estimate of interest, such as substitutional saturation. These considerations can be examined further when genome-scale data become available.

Some processes involved in sexual selection might play an important role in driving the evolution of mitochondrial genes. For example, mitochondria with high rates of energy production are essential for male–male competition but can be detrimental for females, leading to sexually antagonistic selection and rapid mitochondrial evolution [53]. Furthermore, mitochondrial genomes can play a key role in the process of speciation. In species with high mitochondrial mutation rates, allopatric populations can develop mito-nuclear incompatibilities that will tend to accelerate reproductive isolation [54]. Our results suggest that this process is potentially quicker in species with sexual selection, perhaps because sex differences in the optimal expression of mitochondrial genes lead to an evolutionary arms race. In future work, it will be useful to disentangle the effects of nuclear and mitochondrial evolution on the link between sexual selection and diversification.
5. Conclusion

Our study provides evidence that sexual selection and molecular evolution play a joint role in driving rates of diversification in birds. Through its impact on rates of molecular evolution, sexual selection staves off extinction, facilitates the appearance of key innovations and generates reproductive incompatibilities that drive speciation [55,56]. Future research using genome-scale data across many species will help to determine the impact of sexual selection on molecular evolution on a macroevolutionary scale [49]. Molecular data at the genome scale across bird genera will be available within the next few years and will help to connect these micro- and macroevolutionary processes.

To what extent is speciation due to the selection that creates incompatibilities between populations, as opposed to being a by-product of the stochastic fixation of mutations? Strikingly, the results of several studies, including ours, point towards the latter [29]. In addition to the impact of sexual selection on diversification rates being mediated by other life-history traits (e.g. body size in our study), it might also vary across environments (e.g. covery with altitude [57] or ecosystem productivity [36]). Our analyses provide a theoretical and methodological framework to examine the mechanisms by which sexual selection can drive genomic evolution and species diversity patterns that have shaped the tree of life.

Data accessibility. The data used in this paper are available from the Dryad Digital Repository: https://doi.org/10.5061/dryad.p39d12d [58].

Authors' contributions. M.I.-C., M.D.J., S.Y.W.H. and D.A.D. conceived and designed the project; M.I.-C. and D.A.D. collected and analysed the data; all authors contributed to the final version of the manuscript.

Competing interests. We have no competing interests.

Funding. D.A.D., S.Y.W.H. and M.D.J. were supported by the Australian Research Council.

Acknowledgements. We thank Megan Head for comments, and Kristal Cain and Jordi Figuerola for advice on the extraction of data from published sources.

References