Evolution, 2023, **XX(XX)**, 1–13
https://doi.org/10.1093/evolut/qpac010
Original Article

Polygamy and purifying selection in birds

Kees Wanders, PhD[1], Guangji Chen, PhD[2,3], Shaohong Feng, PhD[4,5], Guojie Zhang, PhD[4,5], Tamás Székely, PhD[1,6], Mike Bruford, PhD[7], Zsolt Végvári, PhD[8,9], Götz Eichhorn, PhD[10], Araxi Urrutia, PhD[1,11]

[1]Milner Centre for Evolution, Department of Biology and Biochemistry, University of Bath, Bath, United Kingdom
[2]College of Life Sciences, University of Chinese Academy of Sciences, Beijing, China
[3]BGI-Shenzhen, Shenzhen, China
[4]Liangzhu Laboratory, Zhejiang University School of Medicine, Hangzhou, China
[5]Evolutionary & Organismal Biology Research Center, Zhejiang University School of Medicine, Hangzhou, China
[6]Department of Evolutionary Zoology and Human Biology, University of Debrecen, Debrecen, Hungary
[7]School of Biosciences and Sustainable Places Institute, Cardiff University, Cardiff, United Kingdom
[8]Centre for Ecological Research, Eötvös Loránd Research Network, Institute of Aquatic Ecology, Budapest, Hungary
[9]Senckenberg Deutsches Entomologisches Institut, Müncheberg, Germany
[10]Vogeltrekstation-Dutch Centre for Avian Migration and Demography, Netherlands Institute of Ecology (NIOO-KNAW), Wageningen, Netherlands
[11]Instituto de Ecologia, UNAM, Ciudad de Mexico, Mexico

Corresponding authors: Milner Centre for Evolution, Department of Biology and Biochemistry, University of Bath, Bath, United Kingdom.
Email: kw714@bath.ac.uk; Email: au207@bath.ac.uk

Abstract

Good genes theories of sexual selection predict that polygamy will be associated with more efficient removal of deleterious alleles (purifying selection), due to the alignment of sexual selection with natural selection. On the other hand, runaway selection theories expect no such alignment of natural and sexual selection, and may instead predict less efficient purifying selection in polygamous species due to higher reproductive variance. In an analysis of polymorphism data extracted from 150-bird genome assemblies, we show that polygamous species carry significantly fewer nonsynonymous polymorphisms, relative to synonymous polymorphisms, than monogamous bird species ($p = .0005$). We also show that this effect is independent of effective population size, consistent with the alignment of natural selection with sexual selection and "good genes" theories of sexual selection. Further analyses found no impact of polygamy on genetic diversity, while polygamy in females (polyandry) had a marginally significant impact ($p = .045$). We also recapitulate previous findings that smaller body mass and greater geographic range size are associated with more efficient purifying selection, more intense GC-biased gene conversion, and greater genetic diversity.

Keywords: sexual selection, natural selection, polymorphism, evolutionary genomics, molecular evolution, mating systems

Graphical abstract

Photo by Freya Coursey

Birds exhibit a broad range of mating systems, including monogamous, polyandrous, and polygynous strategies, making them an ideal system to study the evolutionary consequences of mating system (Pitelka et al., 1974). Polygamy has been predicted to influence evolution in a number of ways, primarily due to the association between greater levels of polygamy and more intense sexual selection. In particular, the extent to which sexual selection aligns or interferes with natural selection has been the subject of much debate and remains a controversial area of evolutionary biology (Rowe & Rundle, 2021; Whitlock & Agrawal, 2009).

Elaborate morphological characteristics associated with polygamous mating systems, such as the peacock's tail feathers, are clearly detrimental to individual survival. If the action of sexual selection is restricted to the small subset of genes directly associated with such morphological characteristics, as predicted by runaway selection theories of sexual selection, it will act in opposition to natural selection (Arnold, 1985;

Received May 10, 2022; revisions received September 5, 2022; accepted October 29, 2022
© The Author(s) 2022. Published by Oxford University Press on behalf of The Society for the Study of Evolution (SSE).
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Downloaded from https://academic.oup.com/evolut/advance-article/doi/10.1093/evolut/qpac010/6873148 by guest on 23 January 2023

Fisher, 1958; Kirkpatrick & Ryan, 1991). On the other hand, if sexual selection favors individuals that are healthier in general, as predicted by "good genes" theories of sexual selection, it may act in concert with natural selection to remove harmful alleles and promote adaptation (Agrawal, 2001; Andersson, 1982; Jennions et al., 2001; Siller, 2001; Whitlock & Agrawal, 2009). Sexual selection may also result in an increased mutation rate, due to a trade-off between investment in DNA repair and investment in reproduction (Dowling & Simmons, 2009), selection for rare beneficial mutations when variance in reproductive success is high (Bartosch-Harlid et al., 2003; Petrie, 2021; Petrie & Roberts, 2007), and/or as a result of post-copulatory sperm competition, as greater sperm production requires additional rounds of replication in the male germline (Møller & Cuervo, 2003). Aside from the processes underlying sexual selection, the greater reproductive variance associated with polygamy is expected to reduce the effective size of a polygamous population (Nunney, 1993). Polygamous species are therefore expected to be more affected by genetic drift, which results in less stringent purifying natural selection and reduced genetic diversity (Charlesworth, 2001, 2009; Kimura, 1969a; Wright, 1931). Polygamy is also associated with greater sexual dimorphism, increasing the possibility that alleles harmful to one sex are maintained through their benefit to the other sex (Arnqvist & Rowe, 2013). Finally, recent comparative work in plovers has suggested that polygamous species may exhibit greater gene flow between populations, which could result in an increase in the effective size of a given population (D'Urban Jackson et al., 2017).

Empirical studies of laboratory invertebrate populations have provided evidence for some of these theories, e.g., that the combination of sexual selection and the natural selection improves population fitness relative to natural selection alone (Baur & Berger, 2020; Cally et al., 2019; Jarzebowska & Radwan, 2010; Lumley et al., 2015), and that increasing mate competition can increase mutation rates (Baur & Berger, 2020). However, the question of which processes are most influential remains contentious (Rowe & Rundle, 2021; Whitlock & Agrawal, 2009). Comparative analyses of non-model species can provide insight into this question, and here we analyze the consequences of polygamy on molecular evolution in birds using the largest dataset to date, including single nucleotide polymorphism (SNP) data from 150 species with sequenced genomes. We focus on four hypotheses that make clear predictions for genome-wide signatures of molecular evolution (summarized in Table 1): (1) sexual selection acts in concert with natural selection, by ensuring only the healthiest individuals breed ("good genes" theory of sexual selection), (2) sexual selection is limited to a small number of genes associated with secondary sexual characteristics and preferences and is unrelated to the efficacy of natural selection ("runaway" theory of sexual selection), (3) polygamy acts against natural selection, by lowering the effective population size of a population and thereby increasing the impact of genetic drift, and (4) greater levels of polygamy are associated with a higher mutation rate, either due to a trade-off against DNA repair, selection for a higher mutation rate, or post-copulatory sperm competition. To tease apart the predictions of these hypotheses, we make use of three independent genomic signatures, which reflect the efficiency of purifying selection, the level of genetic diversity, and the intensity of GC-biased gene conversion (gBGC, a fixation bias thought to affect the majority of eukaryotes (Bolívar et al., 2016; Duret & Galtier, 2009; Pessia et al., 2012), although perhaps not

Table 1. Summary of hypotheses linking polygamy and genome-wide molecular evolution, with predictions for the signatures of three evolutionary processes.

Hypothesis	Prediction for purifying selection efficiency	Prediction for neutral genetic diversity	Prediction for GC-biased gene conversion intensity
	(P_n/P_s)†	(Heterozygosity)§	(P_{SW+WS}/P_{SS+WW})‡
(1) Polygamy enhances natural selection via sexual selection (Agrawal, 2001; Andersson, 1982; Jennions et al., 2001; Siller, 2001; Whitlock & Agrawal, 2009)	Polygamous species have more efficient purifying selection (lower P_n/P_s)	No predicted effect	No predicted effect
(2) Polygamy does not enhance natural selection, being limited to the evolution of secondary sexual characteristics and preferences (Arnold, 1985; Fisher, 1958; Kirkpatrick & Ryan, 1991)	No predicted effect	No predicted effect	No predicted effect
(3) Polygamy reduces effective population size (Charlesworth, 2009; Nunney, 1993)	Polygamous species have less efficient purifying selection (higher P_n/P_s)	Polygamous species show reduced heterozygosity	Polygamous species have less intense GC-biased gene conversion (higher P_{SW+WS}/P_{SS+WW})
(4a) Polygamy increases mutation rate via selection for rare beneficial mutations (Bartosch-Harlid et al., 2003; Petrie, 2021; Petrie & Roberts, 2007) or via a trade-off between reproduction and DNA repair (Dowling & Simmons, 2009)	No predicted effect	Polygamous species show greater heterozygosity	No predicted effect
(4b) Polyandry increases mutation rate via sperm competition (Møller & Cuervo, 2003)	No predicted effect	Polyandrous species show greater heterozygosity	No predicted effect

† P_n/P_s = Ratio of GC-conservative nonsynonymous SNPs to GC-conservative synonymous SNPs.
§ Heterozygosity = Proportion of intergenic loci that contain a GC-conservative SNP in a single genome.
‡ P_{SW+WS}/P_{SS+WW} = Ratio of intergenic SNPs affected by GC-biased gene conversion to intergenic SNPs unaffected by GC-biased gene conversion.

Downloaded from https://academic.oup.com/evolut/advance-article/doi/10.1093/evolut/qpac010/6873148 by guest on 23 January 2023

Drosophila (Robinson et al., 2014)). The predictions of each hypothesis for these separate genomic measures are summarized in Table 1.

Materials and methods

Overview of the genomic dataset

Single whole genomes for a total of 150 species were used in this study, including 144 collated as part of the 10,000 bird genomes project (B10k; Feng et al., 2020), and six newly sequenced Arctic shorebird species (*Charadrius hiaticula, Pluvialis squatarola*, *Calidris alpina, Calidris temmincki, Calidris minutus*, and *Phalaropus lobatus*). Species were selected based on the availability of genomes and the relevant life history variables, after excluding flightless birds on the basis that the relationship between geographic range size and body mass and effective population size may be very different in flightless birds, due to reduced constraints on body mass and reduced dispersal ability. The B10k project has deliberately set out to sequence examples from each avian family, and the set of genomes, therefore, includes some particularly long branches leading to families with only one sequenced individual. Long branch lengths cause issues for comparative analyses based on substitutions, such as *dN*/*dS* (the ratio of nonsynonymous to synonymous substitutions) and GC_4 (the GC proportion at fourfold degenerate sites), as differences between species are accumulated along evolutionary periods that might not reflect current phenotypes. This is particularly problematic for fast-evolving behavioral traits such as mating system, for which the entire spectrum of phenotypes can be identified among species of a single family (Pitelka et al., 1974). To avoid these issues, we detect evolutionary signatures in the pattern of polymorphisms, as these reflect more recent evolutionary pressures (McDonald & Kreitman, 1991; Müller et al., 2022). Signatures of genetic diversity and purifying selection efficiency have previously been analyzed using polymorphism data from single genomes (e.g., Figuet et al., 2016), and our analyses of these traits follow established methods: GC-conservative P_n/P_s (the ratio of nonsynonymous to synonymous SNPs) was used for analyzing purifying selection, and intergenic GC-conservative heterozygosity was used for analyzing genetic diversity. In contrast, to our knowledge, previous analyses of gBGC have either relied on substitution data (e.g., Romiguier et al., 2010) or have required multiple genomes with polymorphism data (e.g., Glémin et al., 2015; Muyle et al., 2011; Robinson et al., 2014). Here we present a novel measure of the intensity of gBGC, which makes use of polymorphism data from a single genome: the P_{SW+WS}/P_{SS+WW} ratio. This can be defined as the ratio of intergenic heterozygous sites affected by gBGC to intergenic heterozygous sites unaffected by gBGC.

Explanation and modeling of the P_{SW+WS}/P_{SS+WW} ratio

GC-biased gene conversion (gBGC) results from a meiotic repair bias that favors G and C nucleotides over A and T nucleotides and acts to increase the frequency of "strong" alleles ("S", e.g., G:C) and reduce the frequency of "weak" alleles ("W", e.g., A:T) in a population (Duret & Galtier, 2009; Webster & Hurst, 2012). When a new mutation occurs that introduces a "weak" nucleotide pair in the place of an existing "strong" nucleotide pair (S→W mutation, e.g., G:C→A:T), gBGC reduces the chance of this mutation spreading through the population, analogous to how selection acts on a weakly deleterious allele (Capra et al., 2013; Nagylaki, 1983). However, when a mutation occurs in the opposite direction (W→S, e.g., A:T→G:C), gBGC increases the chance of this mutation spreading through the population, analogous to the effect of selection on a weakly beneficial allele. In contrast, gBGC has no effect on GC conservative mutations, which are rarer mutations that replace "strong" alleles with other "strong" alleles (S→S, e.g., G:C→C:G), or replace "weak" alleles with other "weak" alleles (W→W, e.g., A:T→T:A). GC-biased gene conversion (gBGC) acts in a consistent direction and is expected to be more intense in larger populations (Nagylaki, 1983; Wright, 1931). The effects of gBGC are more pronounced in areas of the genome with high recombination, where the intensity of gBGC is greatest, but nevertheless, they have a significant effect on overall SNP frequencies and genomic GC content (Bolívar et al., 2016). Previous research comparing GC content within the avian clade has found evidence of stronger gBGC in larger populations (Weber et al., 2014), although evidence for this relationship is more mixed in mammals (Kessler & Dean, 2014; Romiguier et al., 2010), and no such relationship has been found across more distantly related animal groups or plants (Clément et al., 2017; Galtier et al., 2018).

Typically, when gBGC strength is measured using polymorphism data, the frequency spectrum of W→S polymorphisms is compared to the frequency of S→W polymorphisms (e.g., Glémin et al., 2015; Muyle et al., 2011; Robinson et al., 2014). However, here we use single whole genomes in a dataset where divergence times between species are often very long, ancestral states cannot be reliably inferred (Hernandez et al., 2007), and S→W and W→S polymorphisms cannot be separated. By modeling the expected heterozygosity levels for the four different SNP categories (S→W, W→S, S→S, and W→W), we show that as long as W→S SNPs are generated more frequently than the reverse, gBGC reduces the total combined number of W→S and S→W SNPs (these can be described as SNPs affected by gBGC). Research into germline mutation rates in eukaryotes has shown consistently that S→W mutations occur more often than the reverse (Bolívar et al., 2016; Hwang & Green, 2004; Lynch, 2010; Ossowski et al., 2010; Smeds et al., 2016; Zhang & Gerstein, 2003) and so the overall frequency of SNPs affected by gBGC is reduced by the action of gBGC. To control for mutation rate differences between species, we divide the total number of S→W and W→S intergenic heterozygous sites with the total number of S→S and W→W intergenic heterozygous sites, to create a measure of gBGC intensity: P_{SW+WS}/P_{SS+WW}. As with all genomic correlates of gBGC intensity, this measure is affected by variation in recombination rates and mutation biases and assumes that such variation is not correlated with the life history traits being compared.

The effect of gBGC on the frequency of W→S and S→W mutations is typically modeled by noting that the rate of gene conversion b is equivalent to a selection coefficient promoting the "strong" allele (e.g., Bólivar et al., 2016; Lartillot, 2013; Mugal et al., 2013). In this approach, W→S mutations are considered weakly beneficial and S→W mutations are considered weakly deleterious, while GC-conservative mutations are neutral. Kimura (1969b) provided equations for estimating the expected amount of heterozygosity in an individual genome for sites under selection (formula 1a), and for selectively neutral sites (formula 1b).

Downloaded from https://academic.oup.com/evolut/advance-article/doi/10.1093/evolut/qpac010/6873148 by guest on 23 January 2023

$$H(p) = \frac{4N_e v_m}{N_e s}\left(\frac{1-e^{-2N_e sp}}{1-e^{-2N_e s}} - p\right) \quad (1a)$$

$$H(p) = 4N_e v_m p(1-p) \quad (1b)$$

Where $H(p)$ = the number of heterozygous sites (per individual), N= the total population size, N_e = the variance effective population size, s = the selection coefficient, v_m = the total number of mutations appearing in the population each generation, and $p = {}^{1}\!/_{2N}$.

By substituting the selection coefficient and mutation rate parameters used by Kimura (1969b) with parameters relevant to S→W, W→S, S→S, and W→W mutations, the effect of population size on the relative proportions of different categories of heterozygous sites can be modeled. These substitutions are summarized in formulas 2a and 2b, where formula 2a applies to mutations affected by gBGC (S→W and W→S) and formula 2b applies to mutations unaffected by gBGC (S→S and W→W). The ratio of polymorphisms affected by gBGC to those unaffected by gBGC (P_{SW+WS}/P_{SS+WW}) is then provided by formula 3.

$$H_{x\to y} = \frac{8N_e^2 \mu_{x\to y} g_x}{N_e b_{x\to y}}\left(\frac{1-e^{-2N_e b_{x\to y} p}}{1-e^{-2N_e b_{x\to y}}} - p\right) \quad (2a)$$

$$H_{x\to x} = 8N_e^2 \mu_{x\to x} g_x p(1-p) \quad (2b)$$

Where x = ancestral nucleotide type (strong or weak), y = derived nucleotide type (strong or weak), μ = mutation rate per site (dependent on x and y), b = gBGC selection coefficient (dependent on x and y), and g = number of sites available for mutation per haploid genome (dependent on x).

$$P_{SW+WS}/P_{SS+WW} = \frac{H_{S\to W} + H_{W\to S}}{H_{S\to S} + H_{W\to W}} \quad (3)$$

Values for the parameters in formulas 2a and 2b were taken from the literature where possible, so that the impact of varying N_e could be modeled in a plausible setting (parameter values summarized in Table 2). Kessler and Dean (2014) noted that N_e estimates in mammals have varied from ~10,000 in humans to ~780,000 in rabbits. Assuming a similar amount of variation in birds, the impact of a 100-fold change in N_e (from 2,000 to 200,000) was modeled, and showed a negative relationship between N_e and P_{SW+WS}/P_{SS+WW} for the full range of intergenic GC content in the genomic dataset (Figure 1). Consistent with these predictions, phylogenetic generalized least squares (PGLS) analysis showed a significant negative correlation between P_{SW+WS}/P_{SS+WW} and intergenic heterozygosity (Table 3). It should be noted that if GC content is sufficiently low, or the mutation rate bias toward S→W is sufficiently weak, so that more W→S mutations are generated than S→W mutations, the predictions of the model are reversed and increasing N_e will increase the predicted P_{SW+WS}/P_{SS+WW} ratio. This switch occurs at GC = ~0.35 for the parameters defined in Table 2.

Figure 1 also highlights a further complexity to the relationship between GC content, $N_e$and P_{SW+WS}/P_{SS+WW}: when N_e is low and gBGC intensity is therefore very weak, both S→W and W→S mutation contribute approximately equally to heterozygosity, and so the greater rate of S→W mutations leads to P_{SW+WS}/P_{SS+WW} increasing in line with the proportion of GC sites. In contrast, when N_e is high and gBGC is having a meaningful impact, S→W mutations are quickly removed and W→S mutations contribute more to heterozygosity, and so an increase in GC sites reduces P_{SW+WS}/P_{SS+WW}. Such an effect did not appear to influence the current analysis, as PGLS models found no interaction between intergenic GC content and heterozygosity, as well as no main effect of intergenic GC content (Table 3). This may reflect the lack of variation in intergenic GC content between species (ranging from 0.4 to 0.44, variance = $3.3e^{-5}$), especially relative to the potentially 100-fold range in effective population size. The complexity of the relationship between P_{SW+WS}/P_{SS+WW}, N_e, and GC content are a limitation of the P_{SW+WS}/P_{SS+WW} measure and may make it unsuitable for certain datasets. Nevertheless, for the current analysis, predictions for the impact of N_e on P_{SW+WS}/P_{SS+WW} are clear, and the measure can provide some insight into the effect of polygamy on molecular evolution.

Table 2. Parameter values used in modeling the formulas 2a and 2b.

Parameter	Value	Justification
g_S	$4.14e^{8}$	1Gb = approx. size of a typical bird genome Average intergenic GC content = 0.414
g_W	$5.86e^{8}$	1Gb – g_S
N_e	50,000	N_e arbitrarily estimated as 50,000 This is smaller than N_e estimated for the mouse, and greater than N_e estimated for the chimp[†]
N	50,000	Equal to N_e to simplify analysis
$\mu_{S\to W}$	$2.51e^{-9}$	Estimated germline mutation rate in the flycatcher[§]
$\mu_{W\to S}$	$1.42e^{-9}$	(as above)
$\mu_{W\to W}$	$2.51e^{-10}$	(as above)
$\mu_{S\to S}$	$4.18e^{-10}$	(as above)
$b_{S\to W}$	$^{-}5e^{-6}$	In mammals, average strength of gBGC $4N_e b =\sim 1$[‡] b estimated as $1/4N_e$ where N_e = 50,000
$b_{W\to S}$	$5e^{-6}$	(as above)

[†] Geraldes et al. (2011); Won and Hey (2005).
[§] Smeds et al. (2016).
[‡] Lartillot (2013).

Downloaded from https://academic.oup.com/evolut/advance-article/doi/10.1093/evolut/qpac010/6873148 by guest on 23 January 2023

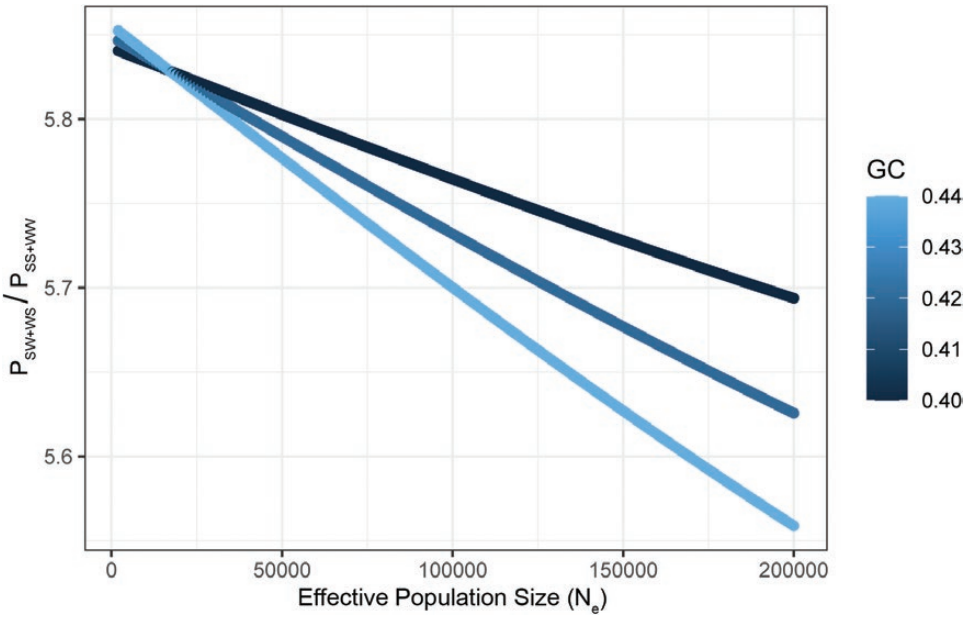

Figure 1. Predicting the impact of N_e (effective population size) on P_{SW+WS}/P_{SS+WW} under plausible parameter values. Predictions are based on calculations using formulas 2a, 2b, and 3, and Table 2 for parameter values.

Table 3. P_{SW+WS}/P_{SS+WW} vs. GC-conservative intergenic heterozygosity + intergenic GC content + genome quality (L50), in a phylogenetic generalized least squares (PGLS) model.[†]

Model term	β (SE) [‡]	*t*-value	*N*	*p*-value
Heterozygosity: GC content	0.046 (0.045)	1.01	150	.31
Heterozygosity	−0.16 (0.042)	−3.82	150	**.0002**
GC content	−0.031 (0.047)	0.65	150	.51
Genome quality (L50)	0.082 (0.042)	1.93	150	.056

[†] PGLS was implemented using Pagel's correlation structure, Pagel's λ = 0.93.
[‡] β = slope (coefficient), *t*-value = slope/standard error, *N* = number of species.

Genomic variables

Single nucleotide polymorphisms (SNPs) were called using GATK (version 3.4-46-gbc02625) (DePristo et al., 2011), and filtered to include only those meeting the following quality criteria: SNPs must be more than 10 bp (base pairs) from another SNP, SNP coverage must be more than one-third mean coverage and less than 2× mean coverage, SNP root-mean-square mapping quality must be at least 25 (as in Nadachowska-Brzyska et al., 2015).

SNPs located within regions of tandem repeats and transposable elements were excluded to avoid the inclusion of spurious SNPs from such regions. Tandem repeats were identified using Tandem Repeats Finder v4.07b41 (Benson, 1999). Transposable elements were identified through homology-based annotation by RepeatMasker (open-4.0.7) with parameters "-nolow -no_is -norna -engine ncbi -parallel 1") at the DNA level based on the Repbase library (v20170127). De novo repeat annotation was completed using RepeatModeler (open-1-0-8) with default parameters to build a de novo repeat library for each assembly (Smit et al., 2015), and this library was also used with RepeatMasker (open-4.0.7) to predict repeats for each species (as in Feng et al., 2020). SNPs located on sex chromosomes were also removed to reduce the noise generated by including a mixture of male and female samples (these SNPs were identified by alignment to chicken sex chromosomes, given the high conservation of synteny among avian species; Ellegren, 2010; Griffin et al., 2007). Locations of SNPs (exonic, intronic, or intergenic), were detected using the protein-coding gene annotation for each species.

The total number of GC-conservative autosomal SNPs passing these quality criteria and located in exons were then identified as synonymous or nonsynonymous, and extracted for analysis of P_n/P_s. Only GC-conservative polymorphisms were included, as gBGC can interfere with signatures of selection (Bolívar et al., 2018). A total of 156 species were initially identified for use in the study, however, three were removed due to a low number of GC-conservative exonic SNPs passing quality control criteria (<200), resulting in a final dataset of 153 species with suitable genomic data. P_n/P_s was calculated by summing the number of GC-conservative nonsynonymous heterozygous sites and dividing this number by 3 times the number of GC-conservative synonymous heterozygous sites (this approximately controls for the greater frequency of new nonsynonymous mutations, as in Figuet et al., 2016). P_n/P_s ratios were natural log (*Ln*)-transformed prior to statistical analysis to reduce the impact of extreme values.

Heterozygosity was calculated for each genome as the number of GC-conservative intergenic SNPs passing quality control criteria, divided by the number of intergenic sites meeting quality control criteria in that genome (as in Figuet et al., 2016). This measure can be defined as the proportion of intergenic sites in a single genome containing GC-conservative SNPs. Heterozygosity was square root transformed before analysis to reduce the impact of extreme values. PGLS analysis revealed no effect of intergenic GC content on heterozygosity ($p > .5$).

To calculate P_{SW+WS}/P_{SS+WW}, autosomal SNPs passing quality criteria and located in intergenic regions were extracted, and the number of SNPs identified as G:T, T:G, G:A, A:G, C:T, T:C, C:A, or A:C was simply divided by the number of SNPs

Downloaded from https://academic.oup.com/evolut/advance-article/doi/10.1093/evolut/qpac010/6873148 by guest on 23 January 2023

identified as A:T, T:A, C:G, or G:C. No transformation was required for this variable.

Life history data

Effective population size is predicted to have a large impact on all the genomic measures analyzed, as purifying selection, gBGC, and genetic diversity are all affected by genetic drift (Charlesworth, 2009). Body mass and geographic range size were therefore included in all models to reduce the unexplained variance, as these variables have previously been found to correlate with population size, and thus may also correlate with effective population size (Damuth, 1981; Gaston & Blackburn, 1996; Greenwood et al., 1996). Body mass estimates were initially collated from the literature by Székely et al. (2022). Where possible, average estimates for males and females were used, but if data was available for just one sex, this was included without adjustment. Body mass was *Ln*-transformed to reduce the impact of extreme values. Distribution ranges were downloaded for all study species as shapefiles from Birdlife.org. Polygons of wintering ranges were then excluded, as these are unrelated to population size when breeding ranges are accounted for. Breeding ranges and year-round resident ranges were retained, and the total geographic range size was calculated using the "areaPolygon" function in the R package "geosphere" (Hijmans, 2012). For all analyses geographic range size was Box-Cox transformed ([geographic range size (km^2)^0.2 – 1]/0.2) to reduce the impact of extreme values. Three of the 153 species with suitable genomic data exhibited outlying phenotypes for geographic range size or body mass (leverage > 2 × [number of variables]/*N*), and these species were excluded from the analysis to avoid spurious associations (as in Thomas, 2015). The final sample size for analyses was therefore 150 species.

For 149 of 150 species, estimates of the extent of polygamy were available from the literature for both sexes, and for the remaining one species *Cuculus canorus*, the extent of polygamy was known for females only (collated by Székely et al., 2022). For the majority of hypotheses outlined in Table 1, the predicted impacts of polygamy in males (polygyny) and polygamy in females (polyandry) are alike, as polygamy in either sex increases the variance in reproductive success and the intensity of sexual selection. Data on the extent of polygyny and polyandry were, therefore, combined for most analyses in order to increase statistical power: species, where >5% of breeding individuals from the more polygamous sex mated multiple times in a season, were considered polygamous (*N* = 29 species), with the rest considered monogamous (*N* = 121 species) (as in D'Urban Jackson et al., 2017). In contrast, the hypothesis that sperm competition increases germline mutation rates predicts an impact of polyandry specifically, as sperm competition is linked to polygamy in females (Cally et al., 2019; Møller, 1991). To test for an effect of sperm competition, heterozygosity was also analyzed in a model comparing polyandrous species (species where >5% of breeding females mate multiple times in a season, *N* = 11) with all other species (*N* = 139 species).

PGLS models showed that there was no significant association between any of the explanatory variables of polygamy, body mass, and geographic range size (Table 4). A separate PGLS model for the 78 species with available census population estimates found that smaller body mass and greater geographic range size were significantly associated with larger census population size (see online supplementary material, Table 1). The lack of a significant correlation between polygamy and census population size suggests there is no severe confounding effect on the dataset, however, only 12 polygamous species had census data available, and so the power to detect an association in this analysis was low. Census population size estimates were taken from three sources—IUCN (2020), BirdlifeInt (2020), and Birds of the World (Billerman et al., 2020), and averages of the extremes were taken when estimates were given as a likely range. This measure was *Ln*-transformed before analysis, to reduce the impact of extreme values (averages of census minimum and maximum estimates were taken after natural log transformation, as these estimates generally followed a logarithmic scale, e.g., "10,000–100,000 individuals").

Software and analysis

All analysis was completed in R version 4.0.1 (R core team, 2020). PGLS analyses were run using the "pgls" function of the caper package, with Pagel's λ estimated by maximum likelihood (Orme et al., 2013). PGLS models were used for all species comparisons, and are a form of linear model that controls for phylogenetic relatedness, in order to avoid issues regarding the nonindependence of data from related species (Symonds & Blomberg, 2014). Statistical assumptions of the models (normality of residuals, no heteroscedasticity) were checked visually by plotting the data, and no issues were detected once variables were appropriately transformed to follow normal distributions, and the three high-leverage species were excluded. Interactions between polygamy/polyandry and the model covariates were checked for in each model, and nonsignificant interactions were removed sequentially to produce the final models (Engqvist, 2005). Body mass and geographic range size were centered and scaled, and the categorical variable of polygamy was also centered, so that main effects could be interpreted in the presence of interactions, and so that slope estimates were comparable among predictor variables (Schielzeth, 2010).

Table 4. Phylogenetic generalized least squares (PGLS) analyses showing life history variable associations. Note that the explanatory variable in each of these pairwise models was selected as the variable with the weakest phylogenetic signal, to avoid conflating the phylogenetic signal with correlation (see online supplementary material, Table 2).

Model	β (SE) †	*t*-value	*N*	*p*-value	Pagel's λ
Polygamy ~ geographic range size	−0.0083 (0.031)	0.27	150	.79	0.54
Body mass ~ polygamy	−0.043 (0.12)	−0.36	150	.72	1.00
Body mass ~ geographic range size	0.082 (0.044)	1.86	150	.064	1

† β = slope (coefficient), *t*-value = slope/standard error, *N* = number of species, λ = phylogenetic signal.

Downloaded from https://academic.oup.com/evolut/advance-article/doi/10.1093/evolut/qpac010/6873148 by guest on 23 January 2023

Genome quality, measured by contig L50, varied widely across species (910–46,581), however, PGLS analysis showed that this measure was not significantly associated with the life history variables studied (see online supplementary material, Table 3). Since a nonsignificant trend toward higher quality genomes in polygamous species was found, all PGLS models involving polygamy were rerun with L50 as a covariate (following square root transformation of L50 to reduce the impact of extreme values). Interpretations from these models were unchanged, suggesting genome quality was not confounding results. Collinearity between independent variables was tested by rerunning all models using the "gls" function of the nlme package (Pinheiro et al., 2017), along with the "corPagel" function of the ape package (Fox et al., 2007), and then applying the "vif" function of the car package (Paradis, 2012). Variance inflation factors for all variables in all models were below 1.3, suggesting a minimal impact of collinearity. Cohen's *D* was calculated using the "cohen.d" function of the "effsize" R package after life history variables were split into binary groups (Torchiano, 2017). Polygamy and polyandry were already binary variables, whereas body mass and geographic range group were simply split around the mean (after the above-mentioned transformation to normal distributions).

Phylogeny

The fourfold-degenerate (4d) site sequences for all 469 1:1 ortholog genes for the initially identified 156 species were used to infer the highest-scoring maximum likelihood tree using a GTRCAT substitution model by RAxML version 8.2.4 (Stamatakis, 2014) and branch lengths were estimated using a GTR substitution model by the phyloFit program in the PHAST package (Siepel & Haussler, 2004).

Data and code availability

Genome sequencing data and genome assemblies of six newly sequenced species generated in this study have been deposited in the NCBI SRA and GenBank (accession PRJNA739535) and CNGBdb (accession CNP0001928). The trait and genomic datasets, as well as all original code, have been deposited at Zenodo, and are publicly available (10.5281/zenodo.7043094). Any additional information required is available from the lead author upon request.

Fieldwork

Blood samples for the six newly sequenced shorebird species (*Charadrius hiaticula, Pluvialis squatarola, Calidris alpina, Calidris temmincki, Calidris minutus*, and *Phalaropus lobatus*), were collected from Kolokolkova Bay (68°35′N, 52°20′E) in Russia. Blood was taken from the brachial vein of adult breeding birds, following established methods that were approved by the University of Bath's Animal Welfare and Ethical Review Body (Székely et al., 2008). No additional permissions were required according to §44 and §6 of the Federal Law of the Russian Federation No. 52 from 24.04.1995 (last update 18.02.2020).

Results and discussion

Purifying selection efficiency

Polygamy may increase purifying selection efficiency due to the alignment of natural and sexual selection (Agrawal, 2001; Andersson, 1982; Jennions et al., 2001; Siller, 2001; Whitlock & Agrawal, 2009), or it may reduce purifying selection efficiency due to a reduction in effective population size (Charlesworth, 2009; Nunney, 1993). A PGLS model analyzing the effect of polygamy, body mass, and geographic range size on purifying selection efficiency (P_n/P_s, the ratio of GC-conservative nonsynonymous to synonymous SNPs), found that polygamous species had a significantly lower P_n/P_s than monogamous species, consistent with polygamy enhancing purifying selection through the alignment of sexual selection and natural selection (Table 5; Figure 2). Significant effects of body mass and range size were also found, which reflect previous studies in a range of taxa showing larger effective population sizes are associated with more efficient purifying selection (Bolívar et al., 2019; Botero-Castro et al., 2017; Corcoran et al., 2017; Figuet et al., 2016; Kutschera et al., 2020; Leroy et al., 2021; Rolland et al., 2020; Romiguier et al., 2014). Effect sizes were calculated independently for each variable using Cohen's *D* (Cohen, 1988). Body mass and polygamy both had a "large" effect size (D = 0.88 and 0.84, respectively), while geographic range size had a "small" effect size (D = 0.15). Variance inflation factor analysis suggested a very weak internal correlation between the covariates (variance inflation factor < 1.3). The relatively small effect of geographic range size may reflect the noise introduced by recent demographic changes in populations, as while polymorphism-based measures of effective population size reflect the average population size of many past generations (Charlesworth, 2009; Müller et al., 2022), current geographic range size will be more closely linked to the current effective population size.

To confirm that the effect of polygamy on purifying selection did not reflect a correlation between polygamy and effective population size, intergenic GC-conservative heterozygosity was added as a covariate to the model (see online supplementary material, Table 4). Comparison of *t*-values following the addition of the heterozygosity covariate revealed greatly reduced explanatory power of body mass (64% reduction in *t*-value) and geographic range size (72% reduction in *t*-value), consistent with effective population size underlying the effect of these variables. However, the explanatory power of polygamy was mostly unaffected (11% reduction in *t*-value), consistent with sexual selection strength underlying the impact of polygamy on purifying selection efficiency. Previous research comparing the efficiency of genome-wide purifying selection with mating systems failed to find this effect (Harrison et al., 2015; Iglesias-Carrasco et al., 2019; Nadeau et al., 2007). The difference in results may reflect a lack of power in the previous studies, resulting from fewer variable genetic sites and/or fewer species. The largest previous study into the question (Iglesias-Carrasco et al., 2019) also included some key

Table 5. Purifying selection efficiency (GC-conservative P_n/P_s) vs. polygamy (presence/absence), geographic range size (km^2), and body mass (g) in a phylogenetic generalized least squares (PGLS) model.

Model term	β (SE) †	*t*-value	*N*	*p*-value
Polygamy	−0.14 (0.038)	−3.57	150	**.0005**
Body mass	0.10 (0.015)	6.70	150	**<.0001**
Geographic range size	−0.039 (0.015)	−2.61	150	**.0099**

† β = slope (coefficient), *t*-value = slope/standard error, *N* = number of species, Pagel's λ = 0.00.

Downloaded from https://academic.oup.com/evolut/advance-article/doi/10.1093/evolut/qpac010/6873148 by guest on 23 January 2023

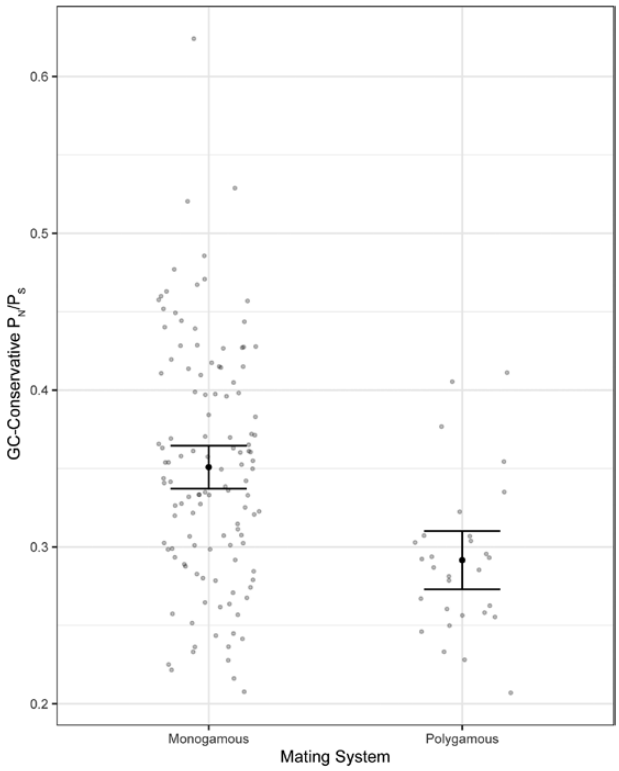

Figure 2. Purifying selection is more efficient in polygamous species (N = 29) than monogamous species (N = 121), as shown by a lower GC-conservative nonsynonymous to synonymous single nucleotide polymorphism (SNP) ratio (P_n/P_s) (p = .0005). Gray dots represent species, black dots represent means, and error bars represent 95% confidence intervals.

methodological differences that may affect the results, such as the use of substitution data (*dN*/*dS*) to measure purifying selection strength, which is more influenced by positive selection than the polymorphism data used here (Smith & Eyre-Walker, 2002), and a focus on polygyny rather than polygamy in general.

Genetic diversity

Various theories have suggested that polygamy, or polyandry specifically, may lead to increased mutation rates and greater genetic diversity (summarized in Table 1). In contrast, genetic diversity may be reduced if polygamous species have smaller effective population sizes (Charlesworth, 2009; Nunney, 1993). Table 6 shows the results of PGLS models analyzing the effect of body mass, geographic range size, and either polygamy or polyandry, on the response variable genetic diversity (intergenic heterozygosity). Significant effects of body mass and geographic range size were found, consistent with many previous studies showing that greater population size is associated with greater genetic diversity (reviewed in Charlesworth, 2009). Body mass had a large effect on heterozygosity (Cohen's D = 1.15), and geographic range size again had a small effect (D = 0.27) (Cohen, 1988). No effect of overall polygamy was found in this model; however, a marginally significant effect of polyandry was detected, with greater genetic diversity in polyandrous species (Figure 3; Cohen's D = 0.67). Greater genetic diversity in polyandrous species is consistent with previous comparative analyses in birds, which have found evidence that higher rates of extra-pair paternity are associated with a greater male mutation

Table 6. GC-conservative intergenic heterozygosity vs. geographic range size (km²), body mass (g), and either polygamy (presence/absence) or polyandry (presence/absence) in a phylogenetic generalized least squares (PGLS) model.

Model term	β (SE) †	*t*-value	*N*	*p*-value
Polygamy	0.0011 (0.0013)	0.87	150	.39
Polyandry	0.0043 (0.0021)	2.02	150	**.045**
Body mass‡	−0.0031 (0.00072)	−4.27	150	**<.0001**
Geographic range size‡	0.0014 (0.00047)	2.98	150	**.0034**

† β = slope (coefficient), *t*-value = slope/standard error, *N* = number of species, Pagel's λ = 0.63.
‡ Results for the body mass and geographic range size variables were almost identical for the two models, and estimates from the polyandry model are presented.

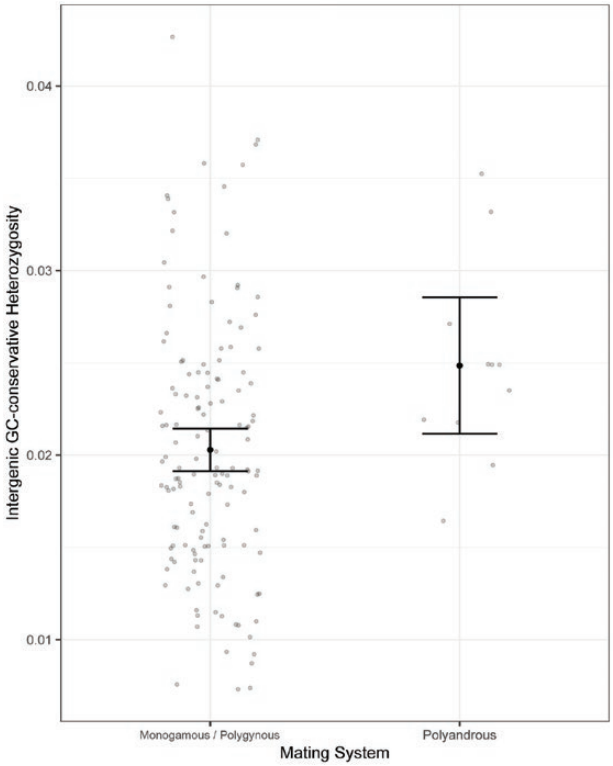

Figure 3. GC-conservative intergenic heterozygosity is higher in polyandrous species (N = 11) than monogamous (and polygynous) (N = 139) (p = .045). Gray dots represent species, black dots represent means, and error bars represent 95% confidence intervals.

bias (Bartosch-Harlid et al., 2003), greater genetic diversity (Gohli et al., 2013; Møller et al., 2008; Petrie et al., 1998), and higher mutation rates (Møller & Cuervo, 2003, 2009), although this was not replicated in swallows (Anmarkrud et al., 2011). Overall, this result provides weak evidence in favor of the hypothesis that sperm competition leads to a greater mutation rate, although it should be treated with caution due to the low number of polyandrous species in the dataset (N = 11 species).

GC-biased gene conversion (gBGC) intensity

The process of gBGC provides an opportunity to understand the impact of polygamy on directional selection without the

Downloaded from https://academic.oup.com/evolut/advance-article/doi/10.1093/evolut/qpac010/6873148 by guest on 23 January 2023

influence of sexual selection, which is not expected to affect gBGC. Polygamy is therefore hypothesized to be associated with reduced gBGC if polygamous species have smaller effective population sizes (Charlesworth, 2009; Nagylaki, 1983; Nunney, 1993) but no other association is predicted (Table 1). Table 7 shows the results of a PGLS analysis comparing the polymorphism-based measure of gBGC, P_{SW+WS}/P_{SS+WW}, with body mass, geographic range size, and polygamy. In this analysis, significant main effects of both body mass and geographic range size were found, consistent with more intense gBGC in larger populations and previous research in birds (Weber et al., 2014). In line with the previous models of purifying selection efficiency and heterozygosity, body mass had a large effect on gBGC (Cohen's D = 1.59), whereas geographic range size had a small effect (D = 0.19) (Cohen, 1988). Whilst no main effect of polygamy was found in this model, a significant interaction with geographic range size was found. Post-hoc PGLS analyses revealed that in polygamous species, gBGC was significantly more intense for species with greater geographic range sizes (p = .005), while in monogamous species, this trend was much weaker and failed to reach significance (p = .22) (Figure 4; see online supplementary material, Table 5). This result may reflect greater gene flow in polygamous species, which would connect disparate parts of a species' range and result in a stronger connection between geographic range size and effective population size (an extension of the "Dispersal to Mate" hypothesis (D'Urban Jackson et al., 2017); illustrated in online supplementary material, Figure 1). However, greater gene flow in polygamous species should also moderate the impact of geographic range size on purifying selection efficiency and heterozygosity, whereas no such interactions were detected (p > .5). It is possible that purifying selection, gBGC, and genetic diversity respond differently to gene flow between populations, however the necessary modeling to make such predictions has not been completed

Table 7. GC-biased gene conversion (gBGC) intensity (intergenic P_{SW+WS}/P_{SS+WW}) vs. polygamy (presence/absence), geographic range size (km^2), body mass (g), in a phylogenetic generalized least squares (PGLS) model, retaining a significant interaction term for polygamy: geographic range size.

Model term	β (SE) †	*t*-value	*N*	*p*-value
Polygamy: geographic range size	−0.30 (0.092)	−3.20	150	**.0017**
Polygamy	−0.021 (0.091)	−0.24	150	.81
Body mass	0.26 (0.062)	4.16	150	**<.0001**
Geographic range size	−0.10 (0.035)	−3.01	150	**.0031**

† β = slope (coefficient), *t*-value = slope/standard error, *N* = number of species, Pagel's λ= 0.95.

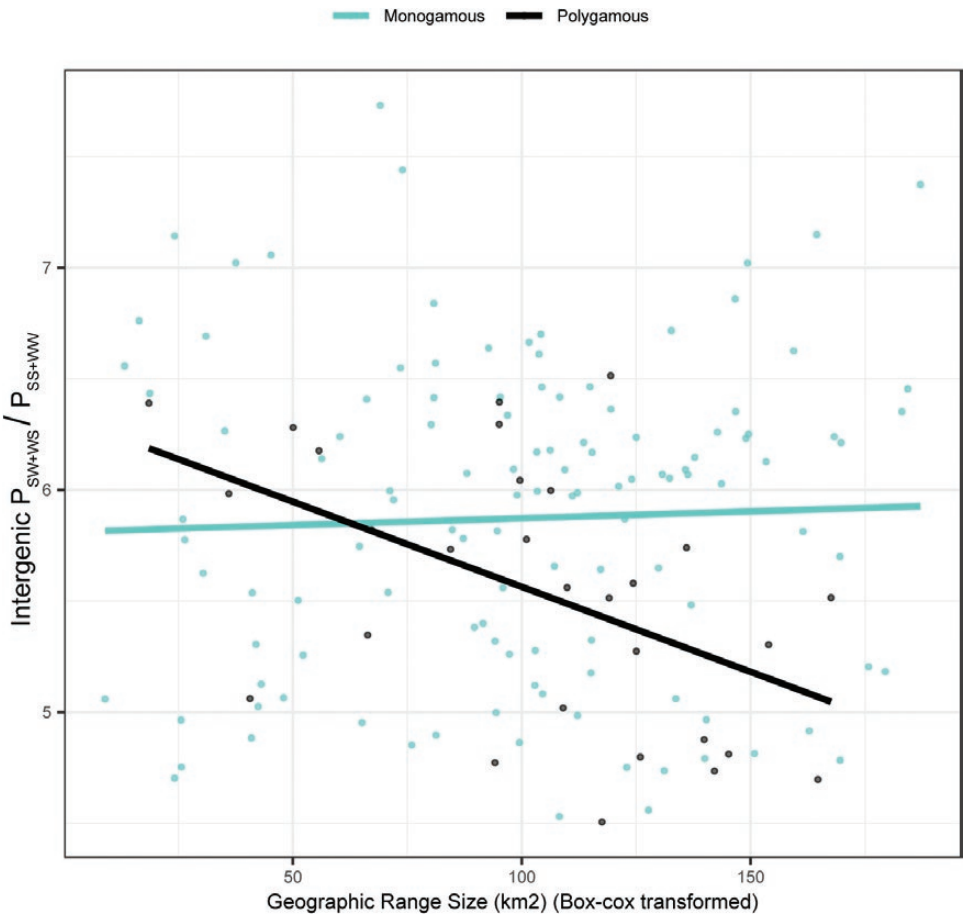

Figure 4. Greater geographic range size is associated with more intense GC-biased gene conversion (gBGC) (p = .0031), measured by the ratio of SNPs affected by gBGC to those unaffected by gBGC (P_{SW+WS}/P_{SS+WW}), and this relationship is stronger for polygamous species (N = 29) than monogamous species (N = 121) (interaction effect: p = .0017). Dots represent species, lines represent linear regressions.

Downloaded from https://academic.oup.com/evolut/advance-article/doi/10.1093/evolut/qpac010/6873148 by guest on 23 January 2023

to our knowledge. A more direct test of the "Dispersal to Mate" theory would compare measurements of gene flow between species, and the interaction between polygamy and geographic range size should be treated with caution until such research is completed.

Summary and conclusions

Polygamy was strongly associated with more efficient purifying selection, whereas no direct effect of polygamy was detected on the signatures of genetic diversity or gBGC intensity (PGLS model results are summarized in online supplementary material, Figure 2). This pattern contrasts with the effects of geographic range size and body mass, which were consistent across all genomic signatures and highlight the large impact of effective population size on genome-wide evolutionary processes. Referring to the predictions in Table 1, the pattern of results for polygamy is consistent with sexual selection enhancing purifying selection for polygamous species, as predicted by "good genes" theories of sexual selection. Wider implications of "good genes" theories of sexual selection include a reduced vulnerability to inbreeding for polygamous species (Jarzebowska & Radwan, 2010), and more efficient adaptation in polygamous species (Lorch et al., 2003), which may in turn underlie the link between sexual selection and diversification (Cally et al., 2021; Iglesias-Carrasco et al., 2019). Furthermore, the pattern of results suggests that the increased variance in reproductive success associated with polygamy does not have a sizeable impact on effective population size, at least relative to the effects of geographic range size and body mass.

A significant effect of polyandry on heterozygosity was also detected, consistent with a mutagenic effect of sperm competition. However, more research is required to corroborate the link between polyandry and genetic diversity, as the current dataset included just 11 polyandrous species. The effect of geographic range size on gBGC intensity was stronger in polygamous species, which is hypothetically consistent with greater gene flow between polygamous populations (D'Urban Jackson et al., 2017), however, the lack of such a moderating effect on purifying selection efficiency or heterozygosity provides evidence against this theory. It should be noted that life history traits and strategies vary widely in the avian class, and while a confounding correlation between effective population size and polygamy was ruled out, it is difficult to exclude the possibility of more complex confounds (e.g., ecological generalism; Tobias & Seddon, 2009). The theories tested in this paper would therefore benefit from further comparative work on a more closely related group of species.

Supplementary material

Supplementary material is available online at *Evolution* (https://academic.oup.com/evolut/qpac010).

Author contributions

K.W. and A.O.U. conceptualized the study. G.C., S.F., G.Z., and K.W. processed and filtered the genome sequence data to extract the dependent variables. T.S., Z.V., and K.W. compiled the phenotypic data. G.C. created the phylogeny used in the analysis. G.E. and K.W. collected the six newly sequenced shorebird samples. K.W. performed all statistical analyses and modeling, created figures and tables, and wrote the article. All authors discussed the results and commented on the manuscript.

Funding statement

This work was supported by the National Environmental Research Council (NE/S007504/1 to KW, NE/P004121/1 to A.O.U.), PAPPIT-DGAPA-UNAM (IA204020 to A.O.U.), Frontiers in Science CONACyT (FORDECYT /24SE/2020/09/30-03—682142 to A.O.U.), the National Natural Science Foundation of China (31901214 and 32170626 to S.F.), the Strategic Priority Research Program of the Chinese Academy of Sciences (XDB31020000 to G.Z.), the International Partnership Program of Chinese Academy of Sciences (152453KYSB20170002 to G.Z.), the Carlsberg Foundation (CF16-0663 to G.Z.), the Villum Foundation (25900 to G.Z.), the Royal Society (WM170050 and APX\R1\191045 to T.S.), the National Research, Development and Innovation Office of Hungary (KKP-126949 to T.S.), the Polar Programme of Netherlands Organization for Scientific Research (ALWPP.2016.030 to G.E.), and by a University of Bath Developing Networks in Europe Grant to A.U. and T.S.

Conflict of interest: The authors have no conflict of interest to declare.

Acknowledgments

Many thanks to Chris Cooney and Alison Wright for their comments on an early version of this manuscript, to Jason Wolf for his comments on the modeling section, and to Chiel Boom and Ingrid Pollet for their help with sample collection.

References

Agrawal, A. F. (2001). Sexual selection and the maintenance of sexual reproduction. *Nature*, 411(6838), 692–695. https://doi.org/10.1038/35079590

Andersson, M. (1982). Sexual selection, natural selection and quality advertisement. *Biological Journal of the Linnean Society*, 17(4), 375–393. https://doi.org/10.1111/j.1095-8312.1982.tb02028.x

Anmarkrud, J. A., Kleven, O., Augustin, J., Bentz, K. H., Blomqvist, D., Fernie, K. J., Magrath, M. J., Pärn, H., Quinn, J. S., Robertson, R. J., Szép, T., Tarof, S., Wagner, R. H., & Lifjeld, J. T. (2011). Factors affecting germline mutations in a hypervariable microsatellite: A comparative analysis of six species of swallows (Aves: Hirundinidae). *Mutation Research*, 708(1–2), 37–43. https://doi.org/10.1016/j.mrfmmm.2011.01.006

Arnold, S. J. (1985). Quantitative genetic models of sexual selection. *Experientia*, 41(10), 1296–1310. https://doi.org/10.1007/BF01952072

Arnqvist, G., & Rowe, L. (2013). *Sexual conflict*. Princeton University Press.

Bartosch-Harlid, A., Berlin, S., Smith, N. G., Mosller, A. P., & Ellegren, H. (2003). Life history and the male mutation bias. *Evolution*, 57(10), 2398–2406. https://doi.org/10.1111/j.0014-3820.2003.tb00251.x

Baur, J., & Berger, D. (2020). Experimental evidence for effects of sexual selection on condition-dependent mutation rates. *Nature Ecology and Evolution*, 4(5), 737–744. https://doi.org/10.1038/s41559-020-1140-7

Benson, G. (1999). Tandem repeats finder: A program to analyze DNA sequences. *Nucleic Acids Research*, 27(2), 573–580. https://doi.org/10.1093/nar/27.2.573

Billerman, S. M., Keeney, B. K., Rodewald, P. G., & Schulenberg, T. S. (2020). *Birds of the world*. https://birdsoftheworld.org/bow/home. Date accessed September 09, 2020.

Downloaded from https://academic.oup.com/evolut/advance-article/doi/10.1093/evolut/qpac010/6873148 by guest on 23 January 2023

BirdLife International. (2020). *IUCN red list for birds*. http://www.birdlife.org. Date accessed September 09, 2020.

Bolívar, P., Guéguen L., Duret, L., Ellegren, H., & Mugal, C. F. (2019). GC-biased gene conversion conceals the prediction of the nearly neutral theory in avian genomes. *Genome Biology*, 20(1), 1–3. https://doi.org/10.1186/s13059-018-1613-z

Bolívar, P., Mugal, C. F., Nater, A., & Ellegren, H. (2016). Recombination rate variation modulates gene sequence evolution mainly via GC-biased gene conversion, not Hill–Robertson interference, in an avian system. *Molecular Biology and Evolution*, 33(1), 216–227. https://doi.org/10.1093/molbev/msv214

Bolívar, P., Mugal, C. F., Rossi, M., Nater, A., Wang, M., Dutoit, L., & Ellegren, H. (2018). Biased inference of selection due to GC-biased gene conversion and the rate of protein evolution in flycatchers when accounting for it. *Molecular Biology and Evolution*, 35(10), 2475–2486. https://doi.org/10.1093/molbev/msy149

Botero-Castro, F., Figuet, E., Tilak, M. K., Nabholz, B., & Galtier, N. (2017). Avian genomes revisited: Hidden genes uncovered and the rates versus traits paradox in birds. *Molecular Biology and Evolution*, 34(12), 3123–3131. https://doi.org/10.1093/molbev/msx236

Cally, J. G., Stuart-Fox, D., & Holman, L. (2019). Meta-analytic evidence that sexual selection improves population fitness. *Nature Communications*, 10(1), 1–10. https://doi.org/10.1038/s41467-019-10074-7

Cally, J. G., Stuart-Fox, D., Holman, L., Dale, J., & Medina, I. (2021). Male-biased sexual selection, but not sexual dichromatism, predicts speciation in birds. *Evolution*, 75(4), 931–944. https://doi.org/10.1111/evo.14183

Capra, J. A., Hubisz, M. J., Kostka, D., Pollard, K. S., & Siepel, A. (2013). A model-based analysis of GC-biased gene conversion in the human and chimpanzee genomes. *PLoS Genetics*, 9(8), e1003684. https://doi.org/10.1371/journal.pgen.1003684

Charlesworth, B. (2001). The effect of life-history and mode of inheritance on neutral genetic variability. *Genetical Research*, 77(2), 153–166. https://doi.org/10.1017/s0016672301004979

Charlesworth, B. (2009). Effective population size and patterns of molecular evolution and variation. *Nature Reviews Genetics*, 10(3), 195–205. https://doi.org/10.1038/nrg2526

Clément, Y., Sarah, G., Holtz, Y., Homa, F., Pointet, S., Contreras, S., Nabholz, B., Sabot, F., Saune, L., Ardisson, M., Bacilieri, R., Besnard, G., Berger, A., Cardi, C., De Bellis, F., Fouet, O., Jourda, C., Khadari, B., Lanaud, C., … Glémin, S. (2017). Evolutionary forces affecting synonymous variations in plant genomes. *PLoS Genetics*, 13(5), e1006799. https://doi.org/10.1371/journal.pgen.1006799

Cohen, J. (1988). *The effect size index: d. Statistical power analysis for the behavioral sciences*. Routledge Academic.

Corcoran, P., Gossmann, T. I., Barton, H. J., Slate, J., & Zeng, K.; Great Tit HapMap ConsortiumGreat Tit HapMap Consortium (2017). Determinants of the efficacy of natural selection on coding and noncoding variability in two passerine species. *Genome Biology and Evolution*, 9(11), 2987–3007. https://doi.org/10.1093/gbe/evx213

D'Urban Jackson, J., Dos Remedios, N., Maher, K. H., Zefania, S., Haig, S., Oyler-McCance S., Blomqvist, D., Burke, T., Bruford, M. W., Székely T., & Küpper C. (2017). Polygamy slows down population divergence in shorebirds. *Evolution*, 71(5), 1313–1326. https://doi.org/10.1111/evo.13212

Damuth, J. (1981). Population density and body size in mammals. *Nature*, 290(5808), 699–700. https://doi.org/10.1038/290699a0

DePristo, M. A., Banks, E., Poplin, R., Garimella, K. V., Maguire, J. R., Hartl, C., Philippakis, A. A., Del Angel, G., Rivas, M. A., Hanna, M., McKenna, A., Fennell, T. J., Kernytsky, A. M., Sivachenko, A. Y., Cibulskis, K., Gabriel, S. B., Altshuler, D., & Daly, M. J. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genetics*, 43(5), 491–498. https://doi.org/10.1038/ng.806

Dowling, D. K., & Simmons, L. W. (2009). Reactive oxygen species as universal constraints in life-history evolution. *Proceedings of the Royal Society B*, 276(1663), 1737–1745. https://doi.org/10.1098/rspb.2008.1791

Duret, L., & Galtier, N. (2009). Biased gene conversion and the evolution of mammalian genomic landscapes. *Annual Review of Genomics and Human Genetics*, 22(10), 285–311. https://doi.org/10.1146/annurev-genom-082908-150001

Ellegren, H. (2010). Evolutionary stasis: The stable chromosomes of birds. *Trends in Ecology and Evolution*, 25(5), 283–291. https://doi.org/10.1016/j.tree.2009.12.004

Engqvist, L. (2005). The mistreatment of covariate interaction terms in linear model analyses of behavioural and evolutionary ecology studies. *Animal Behaviour*, 70(4), 967–971. https://doi.org/10.1016/j.anbehav.2005.01.016

Feng, S., Stiller, J., Deng, Y., Armstrong, J., Fang, Q., Reeve, A. H., Xie, D., Chen, G., Guo, C., Faircloth, B. C., Petersen, B., Wang, Z., Zhou, Q., Diekhans, M., Chen, W., Andreu-Sánchez, S., Margaryan, A., Howard, J. T., Parent, C., … Zhang, G. (2020). Dense sampling of bird diversity increases power of comparative genomics. *Nature*, 587(7833), 252–257. https://doi.org/10.1038/s41586-020-2873-9

Figuet, E., Nabholz, B., Bonneau, M., Mas Carrio, E., Nadachowska-Brzyska, K., Ellegren, H., & Galtier, N. (2016). Life history traits, protein evolution, and the nearly neutral theory in amniotes. *Molecular Biology and Evolution*, 33(6), 1517–1527. https://doi.org/10.1093/molbev/msw033

Fisher, R. A. (1958). *The genetical theory of natural selection*. Рипол Классик.

Fox, J. (2007). *The car package* (pp. 1109). R Foundation for Statistical Computing.

Galtier, N., Roux, C., Rousselle, M., Romiguier, J., Figuet, E., Glémin S., Bierne, N., & Duret, L. (2018). Codon usage bias in animals: Disentangling the effects of natural selection, effective population size, and GC-biased gene conversion. *Molecular Biology and Evolution*, 35(5), 1092–1103. https://doi.org/10.1093/molbev/msy015

Gaston, K. J., & Blackburn, T. M. (1996). Global scale macroecology: Interactions between population size, geographic range size and body size in the Anseriformes. *Journal of Animal Ecology*, 1, 701–714. https://doi.org/10.2307/5669

Geraldes, A., Basset, P., Smith, K. L., & Nachman, M. W. (2011). Higher differentiation among subspecies of the house moutse (Mus musculus) in genomic regions with low recombination. *Molecular Ecology*, 20(22), 4722–4736. https://doi.org/10.1111/j.1365-294X.2011.05285.x

Glémin, S., Arndt, P. F., Messer, P. W., Petrov, D., Galtier, N., & Duret, L. (2015). Quantification of GC-biased gene conversion in the human genome. *Genome Research*, 25(8), 1215–1228. https://doi.org/10.1101/gr.185488.114

Gohli, J., Anmarkrud, J. A., Johnsen, A., Kleven, O., Borge, T., & Lifjeld, J. T. (2013). Female promiscuity is positively associated with neutral and selected genetic diversity in passerine birds. *Evolution*, 67(5), 1406–1419. https://doi.org/10.1111/evo.12045

Greenwood, J. J., Gregory, R. D., Harris, S., Morris, P. A., & Yalden, D. W. (1996). Relations between abundance, body size and species number in British birds and mammals. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 351(1337), 265–278. https://doi.org/10.1098/rstb.1996.0023

Griffin, D. K., Robertson, L. B., Tempest, H. G., & Skinner, B. M. (2007). The evolution of the avian genome as revealed by comparative molecular cytogenetics. *Cytogenetic and Genome Research*, 117(1–4), 64–77. https://doi.org/10.1159/000103166

Harrison, P. W., Wright, A. E., Zimmer, F., Dean, R., Montgomery, S. H., Pointer, M. A., & Mank, J. E. (2015). Sexual selection drives evolution and rapid turnover of male gene expression. *The Proceedings of the National Academy of Sciences*, 112(14), 4393–4398. https://doi.org/10.1073/pnas.150133911

Hernandez, R. D., Williamson, S. H., & Bustamante, C. D. (2007). Context dependence, ancestral misidentification, and spurious signatures of natural selection. *Molecular Biology and Evolution*, 24(8), 1792–1800. https://doi.org/10.1093/molbev/msm108

Hijmans, R. J. (2012). *Introduction to the "geosphere" package (Version 1.5-14)*.

Downloaded from https://academic.oup.com/evolut/advance-article/doi/10.1093/evolut/qpac010/6873148 by guest on 23 January 2023

Hwang, D. G., & Green, P. (2004). Bayesian Markow chain Monte Carlo sequence analysis reveals varying neutral substitution patterns in mammalian evolution. *The Proceedings of the National Academy of Sciences*, 101(39), 13994–14001. https://doi.org/10.1073/pnas.0404142101

Iglesias-Carrasco, M., Jennions, M. D., Ho, S. Y., & Duchêne, D. A. (2019). Sexual selection, body mass and molecular evolution interact to predict diversification in birds. *Proceedings of the Royal Society B*, 286(1899), 20190172. https://doi.org/10.1098/rspb.2019.0172

IUCN. (2020). *The IUCN red list of threatened species*. http://www.iucnredlist.org. Date accessed September 09, 2020.

Jarzebowska, M., & Radwan, J. (2010). Sexual selection counteracts extinction of small populations of the bulb mites. *Evolution; International Journal of Organic Evolution*, 64(5), 1283–1289. https://doi.org/10.1111/j.1558-5646.2009.00905.x

Jennions, M. D., Møller A. P., & Petrie, M. (2001). Sexually selected traits and adult survival: A meta-analysis. *Quarterly Review of Biology*, 76(1), 3–36. https://doi.org/10.1086/393743

Kessler, M. D., & Dean, M. D. (2014). Effective population size does not predict codon usage bias in mammals. *Ecology and Evolution*, 4(20), 3887–3900. https://doi.org/10.1002/ece3.1249

Kimura, M. (1969b). The number of heterozygous nucleotide sites maintained in a finite population due to steady flux of mutations. *Genetics*, 61(4), 893–903. https://doi.org/10.1093/genetics/61.4.893

Kimura, M. (1969a). The rate of molecular evolution considered from the standpoint of population genetics. *The Proceedings of the National Academy of Sciences*, 63(4), 1181–1188. https://doi.org/10.1073/pnas.63.4.1181

Kirkpatrick, M., & Ryan, M. J. (1991). The evolution of mating preferences and the paradox of the lek. *Nature*, 350(6313), 33–38. https://doi.org/10.1038/350033a0

Kutschera, V. E., Poelstra, J. W., Botero-Castro, F., Dussex, N., Gemmell, N. J., Hunt, G. R., Ritchie, M. G., Rutz, C., Wiberg, R. A., & Wolf, J. B. (2020). Purifying selection in corvids is less efficient on islands. *Molecular Biology and Evolution*, 37(2), 469–474. https://doi.org/10.1093/molbev/msz233

Lartillot, N. (2013). Phylogenetic patterns of GC-biased gene conversion in placental mammals and the evolutionary dynamics of recombination landscapes. *Molecular Biology and Evolution*, 30(3), 489–502. https://doi.org/10.1093/molbev/mss239

Leroy, T., Rousselle, M., Tilak, M. K., Caizergues, A. E., Scornavacca, C., Recuerda, M., Fuchs, J., Illera, J. C., De Swardt, D. H., Blanco, G., Thébaud C., Milá, B., & Nabholz, B. (2021). Island songbirds as windows into evolution in small populations. *Current Biology*, 31(6), 1303–1310.e4. https://doi.org/10.1016/j.cub.2020.12.040

Lorch, P. D., Proulx, S., Rowe, L., & Day, T. (2003). Condition-dependent sexual selection can accelerate adaptation. *Evolutionary Ecology Research*, 5(6), 867–881.

Lumley, A. J., Michalczyk, L., Kitson, J. J., Spurgin, L. G., Morrison, C. A., Godwin, J. L., Dickinson, M. E., Martin, O. Y., Emerson, B. C., Chapman, T., & Gage, M. J. (2015). Sexual selection protects against extinction. *Nature*, 522(7557), 470–473. https://doi.org/10.1038/nature14419

Lynch, M. (2010). Rate, molecular spectrum, and consequences of human mutation. *The Proceedings of the National Academy of Sciences*, 107(3), 961–968. https://doi.org/10.1073/pnas.0912629107

McDonald, J. H., & Kreitman, M. (1991). Adaptive protein evolution at the Adh locus in Drosophila. *Nature*, 351(6328), 652–654. https://doi.org/10.1038/351652a0

Møller, A. P. (1991). Sperm competition, sperm depletion, paternal care, and relative testis size in birds. *American Naturalist*, 137(6), 882–906. https://doi.org/10.1086/285199

Møller, A. P., & Cuervo, J. J. (2003). Sexual selection, germline mutation rate and sperm competition. *BMC Evolutionary Biology*, 3(1), 1–11. https://doi.org/10.1186/1471-2148-3-6

Møller, A. P., & Cuervo, J. J. (2009). Minisatellite mutation rates increase with extra-pair paternity among birds. *BMC Evolutionary Biology*, 9(1), 1001–1005. https://doi.org/10.1186/1471-2148-9-100

Møller, A. P., Garamszegi, L. Z., & Spottiswoode, C. N. (2008). Genetic similarity, breeding distribution range and sexual selection. *Journal of Evolutionary Biology*, 21(1), 213–225. https://doi.org/10.1111/j.1420-9101.2007.01450.x

Mugal, C. F., Arndt, P. F., & Ellegren, H. (2013). Twisted signatures of GC-biased gene conversion embedded in an evolutionary stable karyotype. *Molecular Biology and Evolution*, 30(7), 1700–1712. https://doi.org/10.1093/molbev/mst067

Müller, R., Kaj, I., & Mugal, C. F. (2022). A nearly neutral model of molecular signatures of natural selection after change in population size. *Genome Biology and Evolution*, 14(5), evac058. https://doi.org/10.1093/gbe/evac058

Muyle, A., Serres-Giardi, L., Ressayre, A., Escobar, J., & Glémin S. (2011). GC-biased gene conversion and selection affect GC content in the Oryza genus (rice). *Molecular Biology and Evolution*, 28(9), 2695–2706. https://doi.org/10.1093/molbev/msr104

Nadachowska-Brzyska, K., Li, C., Smeds, L., Zhang, G., & Ellegren, H. (2015). Temporal dynamics of avian populations during Pleistocene revealed by whole-genome sequences. *Current Biology*, 25(10), 1375–1380. https://doi.org/10.1016/j.cub.2015.03.047

Nadeau, N. J., Burke, T., & Mundy, N. I. (2007). Evolution of an avian pigmentation gene correlates with a measure of sexual selection. *Proceedings of the Royal Society B*, 274(1620), 1807–1813. https://doi.org/10.1098/rspb.2007.0174

Nagylaki, T. (1983). Evolution of a finite population under gene conversion. *The Proceedings of the National Academy of Sciences*, 80(20), 6278–6281. https://doi.org/10.1073/pnas.80.19.5941

Nunney, L. (1993). The influence of mating system and overlapping generations on effective population size. *Evolution*, 47(5), 1329–1341. https://doi.org/10.1111/j.1558-5646.1993.tb02158.x

Orme, D., Freckleton, R., Thomas, G., Petzoldt, T., Fritz, S., Isaac, N., & Pearse, W. (2013). *Caper: Comparative analyses of phylogenetics and evolution in R (version 0.5. 2)*. https://cran.r-project.org/web/packages/caper/index.html

Ossowski, S., Schneeberger, K., Lucas-Lledó J. I., Warthmann, N., Clark, R. M., Shaw, R. G., Weigel, D., & Lynch, M. (2010). The rate and molecular spectrum of spontaneous mutations in Arabidopsis thaliana. *Science*, 327(5961), 92–94. https://doi.org/10.1126/science.1180677

Paradis, E. (2012). *Analysis of phylogenetics and evolution with R* (Vol. 2). Springer.

Pessia, E., Popa, A., Mousset, S., Rezvoy, C., Duret, L., & Marais, G. A. (2012). Evidence for widespread GC-biased gene conversion in eukaryotes. *Genome Biology and Evolution*, 4(7), 675–682. https://doi.org/10.1093/gbe/evs052

Petrie, M. (2021). Evolution by sexual selection. *Frontiers in Ecology and Evolution*, 9, 786868, 950. https://doi.org/10.3389/fevo.2021.786868

Petrie, M., Doums, C., & Møller A. P. (1998). The degree of extra-pair paternity increases with genetic variability. *The Proceedings of the National Academy of Sciences*, 95(16), 9390–9395. https://doi.org/10.1073/pnas.95.16.9390

Petrie, M., & Roberts, G. (2007). Sexual selection and the evolution of evolvability. *Heredity*, 98(4), 198–205. https://doi.org/10.1038/sj.hdy.6800921

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., Heisterkamp, S., Van Willigen, B., & Maintainer, R. (2017). *Package 'nlme'. Linear and nonlinear mixed effects models, version 3(1)*. https://cran.r-project.org/web/packages/nlme/index.html

Pitelka, F. A., Holmes, R. T., & MacLean, S. F. Jr (1974). Ecology and evolution of social organization in arctic sandpipers. *American Zoologist*, 14(1), 185–204. https://doi.org/10.1093/icb/14.1.185

R Core Team R. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Robinson, M. C., Stone, E. A., & Singh, N. D. (2014). Population genomic analysis reveals no evidence for GC-biased gene conversion in *Drosophila melanogaster*. *Molecular Biology and Evolution*, 31(2), 425–433. https://doi.org/10.1093/molbev/mst220

Downloaded from https://academic.oup.com/evolut/advance-article/doi/10.1093/evolut/qpac010/6873148 by guest on 23 January 2023

Rolland, J., Schluter, D., & Romiguier, J. (2020). Vulnerability to fishing and life history traits correlate with the load of deleterious mutations in teleosts. *Molecular Biology and Evolution*, 37(8), 2192–2196. https://doi.org/10.1093/molbev/msaa067

Romiguier, J., Gayral, P., Ballenghien, M., Bernard, A., Cahais, V., Chenuil, A., Chiari, Y., Dernat, R., Duret, L., Faivre, N., Loire, E., Lourenco, J. M., Nabholz, B., Roux, C., Tsagkogeorga, G., Weber, A. A., Weinert, L. A., Belkhir, K., Bierne, N., … Galtier, N. (2014). Comparative population genomics in animals uncovers the determinants of genetic diversity. *Nature*, 515(7526), 261–263. https://doi.org/10.1038/nature13685

Romiguier, J., Ranwez, V., Douzery, E. J., & Galtier, N. (2010). Contrasting GC-content dynamics across 33 mammalian genomes: Relationship with life-history traits and chromosome sizes. *Genome Research*, 20(8), 1001–1009. https://doi.org/10.1101/gr.104372.109

Rowe, L., & Rundle, H. D. (2021). The alignment of natural and sexual selection. *Annual Review of Ecology, Evolution, and Systematics*, 52(1), 499–517. https://doi.org/10.1146/annurev-ecolsys-012021-033324

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

Siepel, A., & Haussler, D. (2004). Phylogenetic estimation of context-dependent substitution rates by maximum likelihood. *Molecular Biology and Evolution*, 21(3), 468–488. https://doi.org/10.1093/molbev/msh039

Siller, S. (2001). Sexual selection and the maintenance of sex. *Nature*, 411(6838), 689–692. https://doi.org/10.1038/35079578

Smeds, L., Qvarnström A., & Ellegren, H. (2016). Direct estimate of the rate of germline mutation in a bird. *Genome Research*, 26(9), 1211–1218. https://doi.org/10.1101/gr.204669.116

Smit, A. F., Hubley, R., & Green, P. (2015). *RepeatMasker Open-4.0*. http://www.repeatmasker.org

Smith, N. G., & Eyre-Walker, A. (2002). Adaptive protein evolution in Drosophila. *Nature*, 415(6875), 1022–1024. https://doi.org/10.1038/4151022a

Stamatakis, A. (2014). RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30(9), 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Symonds, M. R., & Blomberg, S. P. (2014). *A primer on phylogenetic generalised least squares, in Modern phylogenetic comparative methods and their application in evolutionary biology*. Springer.

Székely, T., Kosztolányi, A., & Küpper, C. (2008). *Practical guide for investigating breeding ecology of Kentish plover*. Charadrius alexandrinus.

Székely, T., Liker, A., Thomas, G. H., Brett, N., Brooks, G., Capp, E., Engel, N., Hodges, S., Hughes, E., Krystalli, A., Lislevand, T., Mapp, A., Pipoly, I., Rice, R., Rossi, L., Komdeur, J., Krüger O., & Gonzalez-Voyer, A. (2022). *Sex roles in birds: Influence of climate, life histories and social environment*. Database: Dryad. https://doi.org/10.5061/dryad.fbg79cnw7

Thomas, R. J. (2015). *Data analysis with R statistical software: A guidebook for scientists*. Eco-explore.

Tobias, J. A., & Seddon, N. (2009). Sexual selection and ecological generalism are correlated in antbirds. *Journal of Evolutionary Biology*, 22(3), 623–636. https://doi.org/10.1111/j.1420-9101.2008.01678.x

Torchiano, M. (2017). *effsize: Efficient effect size computation. R package version 0.7. 1. 1*. https://cran.r-project.org/web/packages/effsize/index.html

Weber, C. C., Boussau, B., Romiguier, J., Jarvis, E. D., & Ellegren, H. (2014). Evidence for GC-biased gene conversion as a driver of between-lineage differences in avian base composition. *Genome Biology*, 15(12), 1–6. https://doi.org/10.1186/s13059-014-0549-1

Webster, M. T., & Hurst, L. D. (2012). Direct and indirect consequences of meiotic recombination: Implications for genome evolution. *Trends in Genetics*, 28(3), 101–109. https://doi.org/10.1016/j.tig.2011.11.002

Whitlock, M. C., & Agrawal, A. F. (2009). Purging the genome with sexual selection: Reducing mutation load through selection on males. *Evolution; International Journal of Organic Evolution*, 63(3), 569–582. https://doi.org/10.1111/j.1558-5646.2008.00558.x

Won, Y. J., & Hey, J. (2005). Divergence population genetics of chimpanzees. *Molecular Biology and Evolution*, 22(2), 297–307. https://doi.org/10.1093/molbev/msi017

Wright, S. (1931). Evolution in Mendelian populations. *Genetics*, 16(2), 97–159. https://doi.org/10.1093/genetics/16.2.97

Zhang, Z., & Gerstein, M. (2003). Patterns of nucleotide substitution, insertion and deletion in the human genome inferred from pseudogenes. *Nucleic Acids Research*, 31(18), 5338–5348. https://doi.org/10.1093/nar/gkg745

Downloaded from https://academic.oup.com/evolut/advance-article/doi/10.1093/evolut/qpac010/6873148 by guest on 23 January 2023