Multivariate selection and intersexual genetic constraints in a wild bird population

Jocelyn Poissant[1,2], Michael B. Morrissey[3], Andrew G. Gosler[4], Jon Slate[2], Ben C. Sheldon[4]

[1]Centre for Ecology and Conservation, College of Life and Environmental Sciences, University of Exeter, Penryn, UK

[2]Department of Animal and Plant Sciences, University of Sheffield, Sheffield, UK

[3]School of Biology, University of St Andrews, St Andrews, UK

[4]Edward Grey Institute, Department of Zoology, University of Oxford, Oxford, UK

Running title: Multivariate intralocus sexual conflict

Correspondence

Jocelyn Poissant

Centre for Ecology and Conservation, College of Life and Environmental Sciences, University of Exeter, Penryn, Cornwall, TR10 9FE, UK. Telephone: +44 (0)1326 370400; Fax: +44 (0)1326 371859; e-mail: j.poissant@exeter.ac.uk

Abstract

When selection differs between the sexes for traits that are genetically correlated between the sexes, there is potential for the effect of selection in one sex to be altered by indirect selection in the other sex, a situation commonly referred to as intralocus sexual conflict (ISC). While potentially common, ISC has rarely been studied in wild populations. Here, we studied ISC over a set of morphological traits (wing length, tarsus length, bill depth, and bill length) in a wild population of great tits (*Parus major*) from Wytham Woods, UK. Specifically, we quantified the microevolutionary impacts of ISC by combining intra- and inter-sex additive genetic (co)variances and sex-specific selection estimates in a multivariate framework. Large genetic correlations between homologous male and female traits combined with evidence for sex-specific multivariate survival selection suggested that ISC could play an appreciable role in the evolution of this population. Together, multivariate sex-specific selection and additive genetic (co)variance for the traits considered accounted for additive genetic variance in fitness was uncorrelated between the sexes (cross-sex genetic correlation = -0.003, 95% CI = -0.83, 0.83). Gender load, defined as the reduction in a population's rate of adaptation due to sex-specific effects, was estimated at 50% (95% CI = 13%, 86%). This study provides novel insights into the evolution of sexual dimorphism in wild populations and illustrates how quantitative genetics and selection analyses can be combined in a multivariate framework to quantify the microevolutionary impacts of ISC.

Keywords: **G** matrix, genetic correlation, intralocus sexual conflict, selection gradient, sexual dimorphism, animal model, heritability, natural selection, quantitative genetics, gender load

Introduction

Males and females in dioecious species are typically dimorphic for a large number of phenotypic traits (Fairbairn *et al.*, 2007). Such sexual dimorphism is generally believed to be adaptive, reflecting difference in sex-specific phenotypic optima (Fairbairn, 2007). While the widespread occurrence of sexual dimorphism indicates that its evolution is possible, large genetic correlations between most homologous male and female traits suggest that its short-term evolution may be constrained (Lande, 1980; Poissant *et al.*, 2010). Indeed, whenever selection differs between the sexes for traits that are genetically correlated between the sexes, there is potential for the effect of selection in one sex to be altered by indirect selection in the other sex, a situation generally referred to as intralocus sexual conflict (ISC) or gender load (Arnqvist & Rowe, 2005; Bedhomme & Chippindale, 2007; Bonduriansky & Chenoweth, 2009; Pennell & Morrow, 2013). While potentially common and important, such intersexual genetic constraints remain little studied in wild populations (Bonduriansky & Chenoweth, 2009; Pennell & Morrow, 2013; Poissant *et al.*, 2010; Wyman *et al.*, 2013).

The evolutionary forces acting on sexual dimorphism depend on the interaction between sex-specific genetic (co)variances and directional selection, as represented by the Lande (1980) sex-specific version of the Lande (1979) equation:

$$\begin{matrix}\Delta z_m \\ \Delta z_f\end{matrix} = \frac{1}{2}\begin{bmatrix}\mathbf{G}_m & \mathbf{B} \\ \mathbf{B}^T & \mathbf{G}_f\end{bmatrix}\begin{bmatrix}\beta_m \\ \beta_f\end{bmatrix} \qquad (1)$$

where Δz_m and Δz_f are vectors of male and female specific responses, $\mathbf{G}_m$ and $\mathbf{G}_f$ are sex-specific additive genetic covariance matrices, $\mathbf{B}$ and $\mathbf{B}^T$ are matrices of cross-sex additive

genetic covariances, and β_m and β_f are sex-specific vectors of selection gradients. The coefficient of one half is included to account for the fact that selected male and female parents make equal autosomal contributions to offspring of both sexes (Lande, 1980). Despite being well known among evolutionary biologists studying sexual dimorphism, surprisingly few have applied the Lande (1980) equation in wild populations (though see Jensen *et al*., 2008, Stearns *et al*., 2012, Tarka *et al*., 2014, and Walling *et al*., 2014, for rare examples). Instead, studies typically focus on estimating either only sex-specific selection (Cox & Calsbeek, 2009) or quantitative genetic parameters (Poissant *et al.*, 2010). In addition, while equation 1 is explicitly multivariate, most quantitative genetic studies of sexual dimorphism performed to date have focused on univariate traits (Wyman *et al.*, 2013). As a consequence, we still know relatively little about the structure of the **B** matrix and its impact on the evolution of sexual dimorphism (Gosden *et al.*, 2012; Wyman *et al.*, 2013). For example, asymmetry of the **B** matrix (differences between below- and above-diagonal elements) may play an important role in facilitating the evolution of multivariate sexual dimorphism (Wyman *et al.*, 2013), but too few **B** matrices have been published to assess the importance of this mechanism (Barker *et al.*, 2010; Wyman *et al.*, 2013). Studies combining sex-specific selection and quantitative genetic parameters, and especially those doing so in a multivariate framework, are therefore needed (Walsh & Blows, 2009; Wyman *et al.*, 2013).

Genetic constraints on the evolution of sexual dimorphism may be widespread (Bonduriansky & Chenoweth, 2009; Cox & Calsbeek, 2009; Pennell & Morrow, 2013; Poissant *et al.*, 2010). In particular, negative cross-sex genetic correlations (r_{mf}) for lifetime fitness in wild populations have been reported (e.g. Brommer *et al*., 2007; Foerster *et al*., 2007), and r_{mf} for fitness components are on average lower than for other trait categories (Poissant *et al.*, 2010). However, little is known about the traits underlying these cross-sex genetic correlations for fitness and their relative importance (Bonduriansky & Chenoweth, 2009; Pennell & Morrow,

104 2013). In part, this is because research tends to be qualitative rather than quantitative, with
105 publications focusing on the statistical significance of intralocus sexual conflicts rather than
106 quantifying their impacts on microevolution.
107
108 A variety of metrics have been developed to quantify multivariate genetic constraints (Walsh &
109 Blows, 2009), and researchers have started applying them to studies of sexual dimorphism in
110 both laboratory (Gosden *et al.*, 2012; Lewis *et al.*, 2011) and wild (Stearns *et al.*, 2012; Tarka *et*
111 *al.*, 2014; Walling *et al.*, 2014) populations. However, in many cases, differences in data
112 transformation and standardization make comparison of results across traits and studies difficult
113 (Hansen & Houle, 2008; Houle *et al.*, 2011). In addition, not all metrics provide easily
114 interpretable or comparable quantitative information (Hansen & Houle, 2008). One approach
115 that is particularly valuable for the study of ISC is the R metric of Agrawal and Stinchcombe
116 (2009). This metric quantifies the impact of genetic covariances on a population's rate of
117 adaptation, including the specific case of cross-sex genetic covariances. Importantly, it yields
118 results that are readily comparable across sets of traits, populations and species (Agrawal &
119 Stinchcombe, 2009). Despite its potential for improving our understanding of ISC, to date few
120 have applied the approach in that context (see Walling *et al*., 2014, for a rare example).
121
122 The importance of considering multivariate phenotypes in studies of ISC is increasingly being
123 recognized (Wyman *et al*., 2013). However, conducting multivariate quantitative genetic studies
124 in wild populations remains challenging, due to difficulties in acquiring sufficiently large
125 pedigree-linked datasets (Wilson & Poissant, 2016). In this study, we take advantage of a long-
126 term study of individual variation in morphological traits (wing length, tarsus length, bill depth
127 and bill length) conducted over multiple decades in a wild pedigreed population of great tits
128 (*Parus major*) from Wytham Woods, Oxford, UK (Savill *et al.*, 2010), to quantify the
129 microevolutionary impacts of ISC in a wild population. Despite being a model organism for

evolutionary ecology research, surprisingly little is currently known about the genetic basis of homologous male and female traits and ISC in this species. This could be due to the fact that morphological traits routinely measured in field studies such as wing and tarsus length are not particularly sexually dimorphic in great tits relative to other bird species (Gosler, 1990; Székely *et al.*, 2007). However, it should be stressed that sexual dimorphism is a relatively poor predictor of contemporary sex-specific selection (Cox & Calsbeek, 2009) and quantitative genetic parameters (Poissant *et al.*, 2010), and hence ISC. In fact, while studies in other great tit populations found little evidence for sex-specific selection on morphology (e.g. Björklund & Linden, 1993), in Wytham Woods, differential use of space and resources by males and females (Gosler, 1987a,b), evidence for sex-specific selection on size (Blakey & Perrins, 1999), and large cross-sex genetic correlations for morphological traits (Garant *et al.*, 2004; Robinson *et al.*, 2013) all suggest that gender load from sex differences in selection on morphology could be substantial. In addition to providing novel insights into the causes and consequences of morphological variation in great tits, this study illustrates some means for generating quantities that will be valuable to quantitatively compare the impacts of ISC over various sets of traits in different populations and species.

Materials and methods

Study population

Great tits are small passerine birds distributed throughout Europe and Asia (Gosler, 1993). Their abundance, wide distribution in Europe, and willingness to use nest boxes, have made them a model of choice in ecology and evolution research, and numerous populations throughout the species' range are now the focus of long-term individual-based studies. The Wytham Woods great tit population has been monitored since 1947. Details about the population and field methods are available in Perrins and Gosler (2010) and references therein. Since 1963, ~1020 nest boxes have been monitored yearly during the breeding season. Each

year, all nestlings (10-15 days post-hatching) and ~80% of presumed parents were captured and fitted with a unique metal ring serving as an ID tag. Additional birds were also captured with nets within and around Wytham Woods as part of specific experiments and long-term monitoring. At each capture, birds were aged and sexed using plumage characteristics and measured for a variety of traits. We assumed that birds not first ringed as nestlings in a Wytham Woods nestbox were immigrants from elsewhere; while a small number of nests probably occur each year in natural cavities these are a small proportion compared to those in nest boxes.

Morphological data

We considered four sexually dimorphic morphological traits that have been consistently measured in adults since 1983: wing length, tarsus length, bill depth and bill length. We considered breeding adults born between 1982 and 2008. For simplicity and to ensure higher repeatability, we limited our analyses to measurements obtained by a single measurer (A. Gosler) who obtained all bill dimension measurements. We only used records of recruits (birds identified attempting reproduction in Wytham Woods) obtained during the nesting season (May and June) in a bird's first year of life. Some individuals (< 0.1%) were measured multiple times and in such cases we used the average. Phenotypic records were available for 2575 individuals measured on average at 3.90 traits each (96.5% of individuals were measured for all traits).

We quantified sexual dimorphism using the size dimorphism index (SDI) of Lovich and Gibbons (1992). It is obtained by subtracting one from the ratio of the larger sex to the smaller sex (i.e. 1 - trait mean of larger sex / trait mean of smaller sex), which sets the neutral value at 0 (i.e. no sexual dimorphism). By convention, values are made positive when female values are the largest and negative when male values are the largest (Lovich & Gibbons, 1992). 95% CI for SDI estimates were obtained by bootstrapping phenotypes 10000 times. We tested if multivariate sexual dimorphism was statistically significant using a MANOVA in R (R Core

Team., 2015).

Pedigree information

A pedigree was constructed based on field information of social parentage from 1958 to 2010. This pedigree included birds ringed within Wytham Woods as well as surrounding woodlands. The pedigree contained 87956 individuals connected by 79400 maternal and paternal links (7187 dams and 7963 sires). Molecular parentage is not routinely conducted in the study population. Given the small number of individuals genotyped relative to the size of the social pedigree and an EPP rate of 12-13% (Firth *et al.*, 2015; Patrick *et al.*, 2012), efforts to combine social and genetic parentage would have affected less than 0.1% of pedigree links, with negligible impacts on quantitative genetic and selection analyses. For simplicity we therefore only used social parentage information. The full social pedigree was used to estimate lifetime reproductive success for selection analyses (details below). For estimating quantitative genetic parameters, we used a trimmed pedigree excluding uninformative individuals generated with the prunePed function in the R package MCMCglmm (Hadfield, 2010). This trimmed pedigree contained 4036 individuals with 1328 unique sires (mean number of offspring per sire ± 1 standard deviation [SD] = 1.83 ± 1.14) and 1313 unique dams (mean number of offspring per dam ± 1 SD = 1.88 ± 1.25), and had a maximum depth of 26 generations.

Quantitative genetic analyses

We partitioned phenotypic variance into additive genetic and other components using a single multivariate animal model and restricted maximum likelihood implemented in ASReml 3.0 (Gilmour *et al.*, 2009). The animal model is a form of mixed model incorporating pedigree information, where the phenotype of each individual is modeled as the sum of its additive genetic value and other random and fixed effects (Kruuk, 2004; Wilson *et al.*, 2010). Fixed

effects, fitted to control for environmental causes of phenotypic resemblance among relatives, included year of birth (fitted as a categorical variable), immigration status (locally raised or not), and information about the environment at each bird's natal nest box (longitude, latitude, altitude and the numbers of oaks within 50 meters; for local birds only). Year of birth was included it as a fixed rather than a random variable to facilitate convergence. Longitude, latitude, altitude and number of oaks within 50 meters were fitted as 4th order polynomials to allow for non-linear relationships. Note that when fixed effects are included trait heritability estimates need to be interpreted as being 'conditioned' on these variables (Wilson, 2008). Mother ID and clutch ID were fitted as random variables in exploratory univariate models but they were generally attributed either little (< 5%) or none of the phenotypic variation. They were therefore not considered in the final multivariate model to facilitate convergence. Ultimately, phenotypic variation after having accounted for fixed effects was therefore partitioned into two components: additive genetic (V_a) and residual (V_r). Inter-sex residual covariances were fixed to zero and genetic correlations were constrained to be between -1 and 1 using the !GZ and !GP arguments in ASReml (Gilmour *et al.*, 2009), respectively. Our choice of starting values for the full multivariate REML was guided by the outputs of simpler models.

Heritability (h^2) was determined by dividing V_a by V_p, where $V_p = V_a + V_r$. To allow comparisons of additive genetic variation among traits and studies (Houle, 1992; Wilson, 2008), we also calculated sex-specific coefficients of variation as

$$CV_a = 100 \times \frac{\sqrt{Va}}{\bar{X}} \qquad (2)$$

and mean-standardized additive genetic variance as

$$I_a = \frac{Va}{\bar{X}^2} \quad (3)$$

Significance of individual additive genetic (co)variance components was tested using likelihood ratio tests. For hypotheses involving parameters on the boundary of parameter space, such as variances, the theoretical asymptotic distribution of the likelihood ratio is a mixture of χ^2 variates, where the mixing probabilities are 0.5, one with 0 degrees of freedom and the other with 1 degree of freedom (Dominicus *et al.*, 2006; Gilmour *et al.*, 2009; Self & Liang, 1987). In these cases, p-values from χ^2 tests with 1 degree of freedom were divided by 2. Likelihood ratio tests were also used to test if individual genetic correlations (r_G) were significantly smaller than one. We tested for significance of variance and covariance estimates using univariate and bivariate models, respectively. To test for multivariate sex × **G** interactions, we compared an unconstrained multivariate model with models where 1) **G** matrices were constrained to be equal between the sexes, 2) genetic variances were constrained to be equal between the sexes, 3) genetic covariances were constrained to be equal between the sexes, and 4) genetic correlations were constrained to be equal between the sexes. This was done using the != argument in ASReml (Gilmour *et al.*, 2009). Because asymmetry of the **B** matrix can play an important role in the evolution of sexual dimorphism (Wyman *et al.*, 2013), we also tested if **B** was asymmetric by comparing an unconstrained model with a model where the corresponding elements from above and below the diagonals of **B** and $\mathbf{B}^T$ were constrained to be equal. Statistical significance was determined using likelihood ratio tests.

Selection analysis

We estimated selection using three fitness metrics. These were the observed number of recruits produced by individuals over their lifetime (lifetime reproductive success, LRS), reproductive longevity (age at last reproduction, hereafter referred to as longevity), and mean annual reproductive success (MRS), calculated as LRS × longevity^{-1}. A recruit was defined as an individual having attempted reproduction in Wytham Woods, and therefore did not include individuals that have only attempted reproduction elsewhere (which is sometimes documented from recapture at other study sites). We restricted selection analyses to individuals that had been measured for all traits simultaneously, and excluded individuals whose nest(s) had been manipulated for experimental purposes such as cross-fostering experiments. Selection coefficients were therefore estimated with fewer records (986 males and 1095 females) than quantitative genetic parameters. Mean observed LRS ± 1 SD was 1.23 ± 1.55 (1.23 ± 1.47 in males and 1.24 ± 1.61 in females). LRS was smaller than the mean number of offspring per parent expected under stable population size (i.e. 2) because a substantial proportion of breeding adults were immigrants, rather than because of a decline in population size. In fact population size has increased over the study period (Garant *et al.*, 2004). Mean longevity was 1.65 ± 1.07 (1.62 ± 1.05 in males and 1.67 ± 1.08 in females), and mean MRS was 0.73 ± 0.87 (0.76 ± 0.90 in males and 0.71 ± 0.84 in females). Variance in relative fitness was 1.57 for LRS (1.43 for males and 1.70 for females), 0.42 for longevity (0.43 for males and 0.42 for females), and 1.40 for MRS (1.38 for males and 1.41 for females).

We tested for the presence of multivariate directional selection using generalized linear models. For LRS, we used a log link function and a negative binomial error structure; for longevity we used a log link function with a poisson error structure; and for MRS, which is a rate, we used a log link function with poisson error while including longevity as weights. In these analyses, sex-specific traits were pooled together after having been centered to sex-specific means of zero.

Significance was tested by comparing models with a fitness component as the dependent variable and no explanatory variable (i.e. only an intercept) with models including all traits as linear explanatory variables. Significance was tested using likelihood ratio tests with 4 degrees of freedom. We then tested for sex × multivariate selection interactions by comparing models with sex and the four traits as linear explanatory variables and models also including all sex × trait interactions (Chenoweth & Blows, 2005).

We estimated sex-specific selection coefficients using the R package GSG version 2.0 (Morrissey & Sakrejda, 2013). Directional (*S*) and quadratic selection differentials were calculated using the moments.differentials function, with standard errors and p-values determined with 10000 bootstraps. Mean-standardized and variance-standardized directional selection differentials were obtained by dividing differentials by trait means and standard deviations, respectively. Quadratic differentials were standardized by dividing by the square of trait means and standard deviations, to obtain mean-standardized and variance-standardized measures, respectively. For this we used trait means and standard deviations obtained from the larger dataset used to estimate quantitative genetic parameters.

We used generalized additive models (GAM) with negative binomial (for LRS) and poisson (for longevity and MRS) error structures fitted using the R package MGCV to identify the most appropriate fitness functions. Initially, we fitted a smooth term (cubic splines) for each trait and all linear interactions. However, when doing so, many smooth terms were penalized to the point of being linear. In that case, meaningful point estimates for quadratic selection gradients could not be obtained (as all curvature of the expected fitness function arises from the curvature of the link function in such instances). Since reporting information about nonsignificant quadratic terms is generally desirable, for example in the context of meta-analyses, we decided to test if there actually was statistical support for fitting smooth terms as opposed to only including linear and

quadratic terms. We did this by comparing models including linear and quadratic predictors with models additionally including smooth terms, with significance of non-linear effects above and beyond quadratic relationships being tested with likelihood ratio tests. Using that approach we found little evidence for non-linear effects above and beyond quadratic relationships, and therefore opted to obtain selection gradients using quadratic models.

Directional (β), quadratic and correlational (γ) selection gradients were obtained using the gam.gradients function of GSG. Standard errors (SE) and p-values for selection gradients were determined with parametric bootstrapping (10000). We obtained mean-standardized and variance-standardized selection gradients (β_u and β_σ) by multiplying directional gradients by trait means and standard deviations, and quadratic and correlational gradients by the square and cross-product of trait means and standard deviations, respectively (Hansen & Houle, 2008). For this we used trait means and standard deviations obtained from the larger dataset used to estimate quantitative genetic parameters. Note that while S_σ is equivalent to obtaining β_σ from a model including a single trait, there is no such direct correspondence between unstandardized and mean-standardized selection differentials and gradients.

As in Stearns *et al.* (2012), we compared the direction of multivariate selection between the sexes by calculating the angle between male and female vectors of directional selection gradients:

$$\theta = \cos^{-1}\left(\frac{a \cdot b}{|a|\,|b|}\right) \qquad (4)$$

where *a* and *b* are the two vectors, $|a| = \sqrt{a \cdot a}$ and $|b| = \sqrt{b \cdot b}$.

An angle of 0° would indicate that multivariate selection is perfectly parallel between the sexes while an angle of 180° would indicate that selection is completely antagonistic. To determine if multivariate selection was significantly parallel or antagonistic (i.e. θ different from the null expectation of 90°) we generated a 95% CI with 10000 sex-specific vectors of selection gradients obtained by parametric bootstrapping in GSG.

Evolutionary responses

The expected responses to selection for sex-specific traits were obtained using Lande's (1980) multivariate equation (equation 1). In order to assess the impact of cross-sex genetic covariances on the evolution of sex-specific traits, we compared predictions from the model above with a model where all elements of the **B** matrix were set to zero. As detailed in Morrissey *et al.* (2012), 95% confidence intervals and standard errors were obtained using 10000 sex-specific vectors of selection gradients generated by parametric bootstrapping in GSG and bootstrap-like replicate **G** matrices by drawing random samples from the sampling variance-covariance matrix of REML estimate of **G**.

Genetic constraints and gender load

The impact of genetic covariances on a population's rate of adaptation can be quantified by comparing the rate of adaptation obtained while considering a full **G** matrix with that obtained while setting all or a subset of genetic covariances to zero (Agrawal & Stinchcombe 2009). We assessed the impact of cross-sex genetic covariances (i.e. the **B** matrix) on the population's rate of adaptation using the R metric of Agrawal and Stinchcombe (2009) while ignoring nonlinear selection (as we are mainly interested in sex-specific directional evolution):

$$R_{\boldsymbol{B}} = \frac{\beta_{mf}{}'\mathbf{G}_{mf}\beta_{mf}}{\beta_{mf}{}'\mathbf{G}_{mf_{(\mathbf{B}=0)}}\beta_{mf}} \quad (5)$$

where is β_{mf} is a vector of sex-specific selection gradients, β_{mf}' is its transpose, $\mathbf{G}_{mf}$ is the additive genetic covariance matrix for sex-specific traits, and $\mathbf{G}_{mf_{(\mathbf{B}=0)}}$ is the **G** matrix where all elements of **B** and $\mathbf{B}^T$ (i.e. cross-sex genetic covariances) are set to zero. A value of $R_{\boldsymbol{B}}$ = 0 would indicate that adaptive evolution of sexual dimorphism is completely precluded by **B**, a value of 1 would indicate that it is not affected by **B,** and values above 1 would indicate that **B** increases adaptive evolution of sexual dimorphism (Agrawal & Stinchcombe, 2009). However, it is important to note that these conclusions are relative to a scenario where traits are not genetically correlated between the sexes. In the absence of any difference in selection between the sexes (i.e. $\beta_m = \beta_f$) and complete overlap of genetic architectures (i.e. $\mathbf{G}_m = \mathbf{G}_f = \mathbf{B}$), $R_{\boldsymbol{B}}$ would take a value of two. We therefore quantified the percent decrease in the population's rate of adaptation due to the presence of separate sexes, or gender load (GL), as

$$\mathrm{GL} = (1 - \frac{R_{\boldsymbol{B}}}{2}) * 100. \quad (6)$$

Note that because we are not considering nonlinear selection, identical R_B and GL values would be obtained when using Hansen and Houle (2008) multivariate evolvability metric instead of Agrawal and Stinchcombe (2009) rate of adaptation (Agrawal & Stinchcombe, 2009).

Genetic variance for fitness implied by selection gradients and G

Evolutionary constraint is any process that reduces the rate of adaptation (increase in mean fitness, or increase (decrease) in a positively (negatively) selected trait, relative to some (presumed) naïve reference rate). Motivated by the fundamental theorem of selection (Fisher 1930), and convincing arguments that constraints should be found in the genetic covariances among traits (Walsh and Blows 2009), the rate of adaptation as represented by some value of the genetic variance of relative fitness, is a particularly useful quantity for evaluating constraint. Any pattern of selection for genetically variable traits implies some genetic variance in relative fitness. For example, in a univariate scenario, the genetic variance in fitness implied by a selection gradient β and an additive genetic variance V_a is $V_{a(w)} = V_a * \beta^2$. Any quantity that reduces this value of $V_{a(w)}$, e.g., selection of a genetically correlated trait, can be seen as a constraint. In the context of studying sexual dimorphism, we can construct a somewhat more subtle measure of constraint due to **B** by calculating sex-specific $V_{a(w)}$ values due only to sex-specific selection and genetic variation, and characterize the extent to which the intersexual genetic covariances in **B** may reduce these values of $V_{a(w)}$.

In the absence of nonlinear selection, the rate of adaptation of Agrawal and Stinchcombe (2009) measures the amount of genetic variance for fitness accounted for by **G** and selection for a set of traits ($\beta'\mathbf{G}\beta$, from formula 12 in Walsh & Blows, 2009). When treating the sexes seperately, population-wide genetic variance in fitness accounted for by sex-specific traits can be obtained by including a factor of ¼ (because we are combining variances; see equation 1 of Wolak *et al*., 2015):

$$V_{a(\beta'\mathbf{G}\beta)} = \frac{1}{4}\beta_{mf}{}'\mathbf{G}_{mf}\beta_{mf} \qquad (7)$$

To obtain sex-specific variances, as well as their covariance, the $\beta_{mf}{}'$ and β_{mf} vectors in equation 7 can be replaced with matrices containing sex-specific selection gradients on different rows, which yields a 2 × 2 sex-specific covariance matrix:

$$V_{a(\beta'\mathbf{G}\beta)} = \frac{1}{4}\begin{bmatrix}\beta_m & 0\\ 0 & \beta_f\end{bmatrix}' \mathbf{G_{mf}} \begin{bmatrix}\beta_m & 0\\ 0 & \beta_f\end{bmatrix} = \frac{1}{4}\begin{bmatrix} V_{a(\beta_m{}'\mathbf{G}_m\beta_m)} & COV_{a(\beta_m{}'\mathbf{G}_m\beta_m,\, \beta_f{}'\mathbf{G}_f\beta_f)} \\ COV_{a(\beta_m{}'\mathbf{G}_m\beta_m,\, \beta_f{}'\mathbf{G}_f\beta_f)} & V_{a(\beta_f{}'\mathbf{G}_f\beta_f)}\end{bmatrix} \quad (8)$$

Population level and sex-specific heritabilities can then be obtained by dividing $V_{a(\beta'\mathbf{G}\beta)}$, $V_{a(\beta_m{}'\mathbf{G}_m\beta_m)}$ and $V_{a(\beta_f{}'\mathbf{G}_f\beta_f)}$ by population-wide, male, and female phenotypic variance in fitness, respectively. Note that when the genetic variance for fitness itself is known, the proportion of the total genetic variation in fitness accounted for by $V_{a(\beta'\mathbf{G}\beta)}$ can also be measured (Walsh & Blows, 2009). However, this was not attempted here because the heritability of fitness in the study population is known to be very small (McCleery *et al.*, 2004). Finally, the standardized cross-sex genetic correlation between sex-specific additive genetic variances in fitness accounted for by the set of traits can be obtained as:

$$r_{mf} = \frac{COV_{a(\beta_m{}'\mathbf{G}_m\beta_m,\, \beta_f{}'\mathbf{G}_f\beta_f)}}{\sqrt{V_{a(\beta_m{}'\mathbf{G}_m\beta_m)} * V_{a(\beta_f{}'\mathbf{G}_f\beta_f)}}}. \quad (9)$$

Results

Phenotypic variation

Multivariate phenotypic sexual dimorphism was statistically significant (MANOVA, $F_{4,2481}$ = 1041.7, P < 0.001). On average, males had longer wings (SDI = -0.039 or 3.9% difference, 95% CI = -0.040, -0.037), longer tarsi (SDI = -0.033, 95% CI = -0.035, -0.031), and deeper (SDI = -0.035, 95% CI = -0.037, -0.032) but shorter bills (SDI = 0.016, 95% CI = 0.014, 0.019) than females.

Quantitative genetic parameters

There was detectable additive genetic variance for all sex-specific traits (Table 1). The proportion of phenotypic variance explained by additive genetic effects after accounting for fixed effects (h^2 ± SE) ranged from 0.53 ± 0.08 for male bill length to 0.78 ± 0.06 for female wing length. In both sexes coefficients of variation (CV_a) and mean-standardized additive genetic variances (I_a) were lowest for wing and tarsus length and highest for bill length and width (Table 1).

Additive genetic covariances were generally positive, and significantly different from zero for approximately half of the trait pairs (Table 2). Genetic correlations (r_G ± SE) within each sex were generally small, with the largest one being between tarsus length and bill depth in males (0.508 ± 0.082). Genetic correlations between the sexes were similarly low, with the exception of cross-sex genetic correlations between homologous traits, which were all large (> 0.8) and not significantly smaller than one.

Male and female **G** matrices were significantly different from each other (Table 2, 2 × ($LogL_1$-$LogL_2$) = 29.38, df = 10, p < 0.01). Genetic variances did not differ significantly between the sexes (2 × ($LogL_1$-$LogL_2$) = 7.62, df = 4, p = 0.11). Genetic covariances and correlations were always smaller in males than in females (Table 2), and these differences were statistically

458 significant (covariances: 2 × ($LogL_1$-$LogL_2$) = 26.16, df = 6, p < 0.001; correlations: 2 × ($LogL_1$-
459 $LogL_2$) = 27.94, df = 6, p < 0.001). The **B** matrix was not significantly asymmetric (2 × ($LogL_1$-
460 $LogL_2$) = 3.36, df = 6, p = 0.76).
461
462 **Selection coefficients**
463 We did not observe significant multivariate directional selection when including all traits from
464 both sexes as explanatory variables in a generalized linear model for either LRS (χ^2 = 4.20, p =
465 0.38), longevity (χ^2 = 5.71, df = 4, p = 0.22), or MRS (χ^2 = 4.89, df = 4, p = 0.30). Similarly, we
466 did not observe significant sex × multivariate selection interaction for LRS (χ^2 = 7.09, df = 4, p =
467 0.13) and MRS (χ^2 = 2.00 , df = 4, p = 0.74). We did, however, observe a significant sex ×
468 multivariate selection interaction for longevity (χ^2 = 14.04, df = 4, p < 0.01).
469
470 Unstandardized, mean, and variance standardized directional selection differentials and
471 gradients for LRS, longevity and MRS are presented in Table 3. Mean standardized directional
472 selection gradients for LRS ranged from -3.103 ± 1.927 for female tarsus length to 3.551 ±
473 1.314 for female bill depth. Only one selection gradient for LRS was statistically significant
474 (female bill depth, β_u = 3.551 ± 1.314, p < 0.01) and this appeared to result mostly from
475 selection through longevity (β_u = 2.661 ± 0.767, p < 0.001). While not statistically significant in
476 males, selection on bill length through longevity was notably different between the sexes (male
477 β_u = 1.473 ± 0.867, p = 0.09, female β_u = -1.493 ± 0.771, p = 0.05).
478
479 The angle between sex-specific vectors of mean-standardized selection gradients for LRS was
480 88.5° (95% CI = 31.5°-147.65°), meaning that multivariate selection in males and females was
481 neither predominantly parallel nor antagonistic. For longevity and MRS, the angle between sex-
482 specific vectors of mean-standardized selection gradients were 128.5° (59.5°, 156.4°) and

50.68° (20.43°, 145.82°), respectively. Overall, selection through longevity was therefore (non-significantly) predominantly antagonistic between the sexes, while selection through MRS was predominantly (non-significantly) parallel.

With the exception of bill depth in males, all point estimates for quadratic selection differentials were negative. However, only those for female bill length were significant different from zero (Appendix S1). No clear tendency emerged for quadratic and correlational selection gradients, with statistical support being generally low (Appendix S2).

Selection responses

Predicted mean-standardized sex-specific responses when including and excluding the **B** matrix are presented in Fig. 1. Point estimates for predictions based on LRS were largest for bill depth and smallest for wing length. Selection through survival was expected to contribute most to the evolution of bill depth, while selection through annual reproductive success was expected to contribute most to the evolution of tarsus length. Patterns appeared to differ between the sexes when setting all elements of **B** to zero. Most notably, bills were predicted to become deeper in females but not in males. However, in general, predicted responses in males and females became nearly identical once including **B**, suggesting little opportunity for the evolution of sexual dimorphism given current multivariate selection and additive genetic (co)variances. Note, however, that 95% confidence intervals generally overlapped between traits, fitness components, and sexes.

Genetic variance for fitness

Together, sex-specific selection through LRS and **G** for the morphological traits considered here accounted for genetic variance explaining less than 1% of the phenotypic variation in relative fitness ($V_{a(\beta'\mathbf{G}\beta)}$ = 0.0025, 95% CI = 0.0020, 0.0137; h^2 = 0.0016, 95% CI = 0.0013, 0.0086,

509 Table 4). About 2/3 of this genetic variance was related to selection through MRS ($V_{a(\beta'\mathbf{G}\beta)}$ =
510 0.0016, 95% CI = 0.0004, 0.0084), while the reminder was related selection through longevity
511 ($V_{a(\beta'\mathbf{G}\beta)}$ = 0.0008, 95% CI = 0.0003, 0.0036). Sex-specific estimates are presented in Table 4.
512
513 The correlation between male and female genetic variance for relative fitness accounted for by
514 the set of morphological traits was -0.003 (95% CI = -0.83, 0.83). Longevity and MRS, when
515 considered in isolation, accounted for genetic variances in relative fitness that were negatively -
516 0.43 (95% CI = -0.86, 0.58) and positively 0.59 (95% CI = -0.78, 0.91) correlated between the
517 sexes, respectively. The ratio of $V_{a(\beta'\mathbf{G}\beta)}$ obtained while including **B** to $V_{a(\beta'\mathbf{G}\beta)}$ obtained while
518 excluding **B** was 1 ($R_{\mathbf{B}}$ = 1.00, 95% CI = 0.27 – 1.73). Cross-sex genetic covariance therefore
519 did not impact $V_{a(\beta'\mathbf{G}\beta)}$ relative to a situation where traits were not genetically correlated
520 between the sexes. On the other hand, the presence of separate sexes, relative to a situation
521 where there would be no differences in selection and genetic architectures between the sexes,
522 resulted in a gender load of 50% (95% CI = 13, 86). Gender load estimates for longevity and
523 MRS were 68 % (95% CI = 25, 90) and 26% (95% CI = 8, 84), respectively.
524
525

526 **Discussion**

527 Significant additive genetic variance was detected for all traits, indicating that responses to
528 selection and genetic constraints were possible. Corresponding heritability estimates were
529 large, as is usually the case for morphological traits in birds (Merilä & Sheldon, 2001) including
530 previous estimates in Wytham Woods great tits obtained using a variety of methods (Gosler,
531 1987a; Robinson *et al.*, 2013; Santure *et al.*, 2015). Coefficients of variation (CV_a) were also
532 typical of morphological traits in other species (Houle, 1992).
533

534 **G** matrices differed between the sexes, with covariances (and genetic correlations) being
535 consistently smaller in males than in females. Sex differences in **G** are relatively common and
536 have, for example, been documented in a number of vertebrates (Arnold & Phillips, 1999;
537 Jensen *et al.*, 2003), invertebrates (Lewis *et al.*, 2011; Rolff *et al.*, 2005), and plants (Ashman,
538 2003; Campbell *et al.*, 2011; McDaniel, 2005; Steven *et al.*, 2007). Such differences are
539 important because they indicate that the sexes could respond differently to direct and indirect
540 selection. The larger genetic covariances in females suggest that genetic integration of
541 morphological traits may be greater in that sex. The reasons why that would be are unclear but
542 one possibility could be the presence of sex differences in correlational selection (McGlothlin *et*
543 *al.*, 2005). Extra-pair paternities (EPP) may also have contributed to these patterns, a point we
544 return to below.
545
546 The evolution of sexual dimorphism depends on the structure of the **B** matrix, which includes
547 genetic covariance between homologous as well as non-homologous male and female traits
548 (Lande, 1980; Wyman *et al.*, 2013). Genetic correlations between homologous male and female
549 traits were all very large, which was similar to previous findings for wing length and fledgling
550 mass in the same population (Garant *et al.*, 2004; Robinson *et al.*, 2013). Large genetic
551 correlations for traits exhibiting relatively low level of sexual dimorphism was consistent with the
552 tendency for cross-sex genetic correlations and sexual dimorphism to be negatively correlated
553 (Poissant *et al.*, 2010). Combined with an absence of significant differences in additive genetic
554 variance between the sexes, our results suggests that the short-term evolution of sexual
555 dimorphism in Wytham Woods great tits may be limited for many aspects of morphology
556 (Lande, 1980). In contrast, genetic correlations between non-homologous traits were
557 comparatively small and at first sight appeared to play a smaller role in constraining the
558 evolution of sexual dimorphism; although assessing the constraining effect of individual genetic
559 correlations can be misleading (Walsh & Blows, 2009). Finally, for the traits considered here,

asymmetry of the **B** matrix (i.e. of its off-diagonal elements, Wyman *et al*., 2013) did not appear to play a role in facilitating the evolution of sexual dimorphism.

The presence of a significant sex × multivariate selection interaction for longevity indicated that aspects of morphology, or correlated traits, were under sex-specific directional survival selection. However, this pattern was attenuated and no longer statistically significant once combined with variation in MRS (i.e. when considering LRS). This illustrates how considering various fitness components can increase knowledge about the biology of selection and constraints, but also how individual fitness components, when treated in isolation, may lead to erroneous evolutionary predictions. In the context of ISC, it also stresses out the need to interpret and compare studies in the context of the fitness component used. For example, Tarka *et al*. (2014) also studied ISC over morphological traits in a wild bird population using LRS but they defined LRS as the total number of fledglings produced over an individual's lifetime whereas we defined LRS as the total number of recruits produced. While results from the two studies are similar, they are therefore not entirely equivalent because the LRS metric used by Tarka *et al*. (2014) did not include selection through survival to adulthood and sexual selection (i.e. finding a mate) whereas the one used in herein did.

We detected significant directional selection for female bill depth when considering LRS, and this pattern appeared to result primarily from viability selection. Bill morphology is a classic example of a selected trait in birds, as is it closely tied to variation in the availability of different food types. The strength of selection for female bill depth was relatively strong, as a β_u of 3.55 is larger than the 75% percentile for β_u in natural populations (β_u = 1.34) compiled by Hereford *et al*. (2004). The selection gradient for female bill depth was also especially large considering that our sample size was greater than most published studies to date and that large sample sizes tend to yield smaller, more accurate, estimates (Hereford *et al*., 2004). In contrast, bill depth did

not appear to be under directional selection in males. Bill length, another important aspect of bill morphology, appeared to be under sexually antagonistic viability (longevity) selection, but this pattern was not mirrored by selection through MRS. As a consequence, evidence for sexually antagonistic selection on bill length was attenuated when considering selection through LRS. Sex differences in selection on bill morphology, believed to arise from sex differences in food utilization, have been documented in other systems. For example, in a wild population of serin (*Serinus serinus*), survival selection on bill morphology was directional in females but stabilizing in males (Björklund & Senar, 2001). Male and female great tits are known to exploit different dietary niches in Wytham Woods (Gosler, 1987a,b) and this could explain patterns documented herein. Additional research on the drivers of sex-specific selection on bill morphology and associated genetic constraints would be valuable; for example on the impact of spatial and temporal heterogeneity in food availability and niche partitioning.

Extra-pair paternity (EPP) has been estimated at 12-13% in the study population (Firth *et al.*, 2015; Patrick *et al.*, 2012) and these could have affected selection coefficients and quantitative genetic parameters estimates. This situation is similar to other studies where molecular parentage analyses are not routinely conducted, such as in humans (e.g. Bolund *et al*., 2013; Stearns *et al*., 2012). EPP introduce errors in male LRS estimates, which may unduly reduce covariance between LRS and trait variation in that sex. EPP is also expected to limit phenotypic resemblance between offspring and their (social) father as well as other relatives (e.g. paternal grand-parents), which could reduce additive genetic variance and heritability of both male and female traits estimated from an animal model but more so for male traits (Brommer *et al.*, 2005; Brommer *et al.*, 2007; Charmantier & Réale, 2005; Jensen *et al.*, 2003; Morrissey *et al.*, 2007). Reduced phenotypic resemblance between offspring and paternal relatives could also reduce genetic covariance within and between the sexes. We would expect such a bias to be most pronounced for male-specific genetic covariances, followed by cross-sex and female-specific

612 covariances. Larger covariances in females compared to males were consistent with this
613 predicted pattern. However, it is worth noting that sex differences in genetic correlations were
614 substantial, and that Morrissey *et al.* (2007) found that, for the most part, genetic correlations
615 are usually unbiased by pedigree errors because covariances and variances are usually
616 underestimated in similar proportions. The large differences between male and female genetic
617 correlations therefore suggest that sex differences in quantitative genetic parameters were
618 unlikely due to EPP alone. Nonetheless, the potential for EPP to bias estimates means that any
619 downstream sex differences in evolutionary predictions should be interpreted with caution.
620
621 In this study we have quantified the evolutionary consequences of ISC over a set of
622 morphological traits in a population of great tits by estimating the impacts of sex-specific
623 selection and genetic variance on the population's rate of adaptation. At face value, a gender
624 load of 50% for a set of traits exhibiting little sexual dimorphism appeared substantial. In
625 comparison, in a similar study in Red Deer (*Cervus elaphus*) by Walling *et al*. (2014), gender
626 load for a set of life history traits was estimated at 27.5% (calculated from their multivariate
627 evolvability ratio of 1.45). However, additional studies where a similar approach is applied will
628 be needed to reach conclusions on the relative importance of ISC quantified here and by
629 Walling *et al*. (2014). Estimates were also arguably imprecise, but since the current study and
630 the one of Walling *et al*. (2014) were based on two of the world's largest datasets for wild
631 pedigreed populations, similarly or even less precise results are to be expected as researchers
632 work toward quantifying the impacts of ISC in other systems. This is perhaps not surprising
633 given that the estimation of genetic covariances is known to require large sample sizes (Lynch,
634 1999) and that selection analyses in wild populations are often underpowered (Hersch &
635 Phillips, 2004). In that context, the joint publication of **B** matrices and sex-specific selection
636 gradients should be encouraged, even in the absence of significant results, as compiling results
637 from a large number of studies will be necessary to contextualize results and gain a broader

638 understanding of the importance of ISC in constraining contemporary evolution in natural
639 populations (Cox & Calsbeek, 2009; Poissant *et al.*, 2010; Wyman *et al.*, 2013).
640
641 **Acknowledgments**
642 We thank the many people and funding sources, too numerous to list here, that have
643 contributed to the Wytham Woods great tit project over the years. We would also like to thank
644 Anna Santure, Mark Adams and Matt Robinson for constructive discussions during the
645 formulation and execution of this work. Matthew Wolak and anonymous referees provided
646 valuable feedback on previous versions of this manuscript. This work was part of JP's post-
647 doctoral research at the University of Sheffield funded by the Natural Sciences and Engineering
648 Council of Canada (NSERC) and a Marie Curie International Incoming Fellowship. JP was
649 supported by a Leverhulme Trust Early Career Fellowship. MBM was supported by a Royal
650 Society University Research Fellowship.
651
652
653
654
655
656
657
658
659
660
661
662
663

Table 1. Number of individuals, raw trait means (in millimetres), sexual dimorphism index (SDI) and univariate quantitative genetic parameters for sex-specific morphological traits in a wild population of great tits. Phenotypic variances after having accounted for fixed effects (V_p) and additive genetic variances (V_a) were estimated using a multivariate animal model. Heritability ($h^2 = V_a / V_p$), coefficient of variation (CV_a) and mean-standardized additive genetic variance (I_a) are also presented. Standard errors are presented in parentheses. Statistical significance of V_a was tested using likelihood ratio tests.

trait	n	mean (sd)	SDI	V_p	V_a	h^2	CV_a	Ia * 10000
<u>males</u>								
wing length	1207	75.86 (1.32)	-0.039	1.663 (0.072)	1.255 (0.132)***	0.75 (0.06)	1.48 (0.08)	2.18 (0.23)
tarsus length	1171	23.71 (0.52)	-0.033	0.259 (0.012)	0.164 (0.022)***	0.63 (0.07)	1.71 (0.11)	2.91 (0.38)
bill depth	1167	4.59 (0.14)	-0.035	0.017 (0.001)	0.010 (0.001)***	0.57 (0.07)	2.15 (0.16)	4.63 (0.69)
bill length	1167	13.46 (0.41)	0.016	0.156 (0.007)	0.082 (0.013)***	0.53 (0.08)	2.13 (0.17)	4.54 (0.74)
<u>females</u>								
wing length	1367	73.05 (1.30)	-	1.603 (0.066)	1.247 (0.121)***	0.78 (0.06)	1.53 (0.07)	2.34 (0.23)
tarsus length	1321	22.95 (0.53)	-	0.264 (0.011)	0.204 (0.022)***	0.77 (0.06)	1.97 (0.11)	3.87 (0.41)
bill depth	1322	4.44 (0.14)	-	0.018 (0.001)	0.011 (0.001)***	0.64 (0.07)	2.40 (0.16)	5.78 (0.75)
bill length	1322	13.68 (0.47)	-	0.222 (0.009)	0.136 (0.018)***	0.61 (0.07)	2.69 (0.18)	7.25 (0.98)

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.