Sex-specific additive genetic variances and correlations for fitness in a song sparrow (*Melospiza melodia*) population subject to natural immigration and inbreeding

Abstract

Quantifying sex-specific additive genetic variance (V_A) in fitness, and the cross-sex genetic correlation (r_A), is prerequisite to predicting evolutionary dynamics and the magnitude of sexual conflict. Further, quantifying V_A and r_A in underlying fitness components, and genetic consequences of immigration and resulting gene flow, is required to identify mechanisms that maintain V_A in fitness. However, these key parameters have rarely been estimated in wild populations experiencing natural environmental variation and immigration. We used comprehensive pedigree and life history data from song sparrows (*Melospiza melodia*) to estimate V_A and r_A in sex-specific fitness and underlying fitness components, and to estimate additive genetic effects of immigrants alongside inbreeding depression. We found evidence of substantial V_A in female and male fitness, with a moderate positive cross-sex r_A. There was also substantial V_A in male but not female adult reproductive success, and moderate V_A in juvenile survival but not adult annual survival. Immigrants introduced alleles with negative additive genetic effects on local fitness, potentially reducing population mean fitness through migration load, but alleviating expression of inbreeding depression. Our results show that V_A for fitness can be maintained in the wild, and be broadly concordant between the sexes despite marked sex-specific V_A in reproductive success.

KEY WORDS: Cross-sex genetic correlation, genetic groups, inbreeding depression, migration load, quantitative genetic generalized linear mixed model, sexual conflict

Introduction

The magnitude of additive genetic variance (V_A) in fitness governs the rate of adaptive trait evolution and the expected increase in population mean fitness (Fisher 1930; Robertson 1966; Price 1970), and thereby links adaptation and population persistence (Bell 2013; Gomulkiewicz and Shaw 2013; Carlson et al. 2014; Shaw and Shaw 2014). Quantifying the magnitude of V_A in fitness, and identifying key mechanisms that maintain or constrain such V_A, are consequently central objectives in evolutionary biology (Burt 1995; Barton and Keightley 2002; Ellegren and Sheldon 2008; Walsh and Blows 2009; Shaw and Shaw 2014; Hendry et al. 2018).

Fitness can be defined and measured in numerous ways (Brommer 2000; Metcalf and Pavard 2007; Orr 2009; Sæther and Engen 2015). In the context of Fisher's (1930) Fundamental Theorem, absolute fitness is most straightforwardly defined as the total number of zygotes produced by a zygote (Crow and Kimura 1970; Arnold and Wade 1984; Falconer 1989 p. 336; Shaw and Shaw 2014). Such fitness emerges from a sequence of components comprising survival from conception to sexual maturity and adult lifetime reproductive success (LRS). Adult LRS itself results from a repeating sequence of reproduction and survival to the next reproductive opportunity, eventually terminated by death. Therefore, the magnitude and maintenance of V_A in fitness will ultimately depend on the magnitudes of V_A in all fitness components, and on the additive genetic correlations (r_A) among these components.

In organisms with separate sexes, many genes that affect fitness are expressed in both sexes and can have congruent or divergent pleiotropic effects on sex-specific fitness components (Arnold and Wade 1984; Falconer 1989, p.338; Chippindale et al. 2001). Resulting r_As between the sexes, and among fitness components within each sex, can generate evolutionary sexual conflict, and multiple life history trade-offs and multi-dimensional constraints (Lande 1980, 1982; Rose 1982; Charlesworth 1987; Chippindale et al. 2001; Kruuk et al. 2008; Bonduriansky and Chenoweth 2009; Walsh and Blows 2009; Shaw and Shaw 2014). Consequently, V_A in sex-specific fitness and

fitness components, and corresponding cross-sex and within-sex r_As, are key parameters shaping the total V_A for fitness that emerges and is maintained following selection (Lewontin 1974; Rose 1982; Chippindale et al. 2001; Brommer et al. 2007; Kruuk et al. 2008; Walsh and Blows 2009; Walling et al. 2014).

Further, the magnitude of standing V_A in fitness in any focal subpopulation will also depend on natural spatio-temporal variation in the form of selection and local adaptation and associated patterns of immigration and inter-deme gene flow (Merilä and Sheldon 1999; Zhang 2012; Carlson et al. 2014; Shaw and Shaw 2014). Immigration could increase V_A by introducing alleles with negative or positive additive effects on local fitness, potentially causing migration load, and impeding or facilitating adaptation and population growth (Lenormand 2002; Garant et al. 2007; Edelaar and Bolnick 2012; Carlson et al. 2014). Immigration could further change mean fitness by altering the degree of local inbreeding versus outbreeding and associated expression of inbreeding depression, heterosis, and outbreeding depression (Ingvarsson and Whitlock 2000; Tallmon et al. 2004; Frankham 2016). Such effects, and resulting effective gene flow, depend fundamentally on the genetic properties of immigrants relative to focal natives (Ingvarsson and Whitlock 2000; Tallmon et al. 2004; Edelaar and Bolnick 2012). Therefore, understanding and predicting overall evolutionary dynamics not only requires estimation of V_A in fitness and underlying fitness components in both sexes, and associated cross-sex and within-sex r_As (Ellegren and Sheldon 2008; Kirkpatrick 2009; Kruuk et al. 2008, 2014; Shaw and Shaw 2014; Walling et al. 2014), but also requires explicit estimation of multiple genetic effects resulting from immigration (Ingvarsson and Whitlock 2000; Lenormand 2002; Tallmon et al. 2004; Garant et al. 2007; Edelaar and Bolnick 2012; Carlson et al. 2014).

Fitness and its components reflect the expression of numerous developmental, physiological, morphological, and behavioral traits, and are consequently best conceptualized as highly polygenic, complex traits (e.g. Houle 1992; Barton and Keightley 2002; Flint and Mackay 2009; Hill 2012;

Travisano and Shaw 2013), although loci of large effect can exist (e.g. Johnston et al. 2013; Trask et al. 2016). Hence, key V_As and r_As can be estimated using quantitative genetic methods derived from the infinitesimal model (Lynch and Walsh 1998). Although the phenotypic distribution of fitness is intrinsically non-Gaussian (Arnold and Wade 1984; Wagenius et al. 2010; Shaw and Etterson 2012; Bell 2013; Shaw and Shaw 2014), V_As and r_As can be estimated on latent scales. Here, non-Gaussian phenotypic expression is assumed to reflect underlying variation in a normally distributed latent trait to fulfill the fundamental quantitative genetic assumption of multivariate normality of the average effect of an individual's polygenic genotype (i.e. breeding value, Lynch and Walsh 1998 pp.72-79 and Ch.25; de Villemereuil et al. 2016).

Estimating V_As and r_As in latent traits in wild populations is empowered by a class of quantitative genetic generalized linear mixed models (QGGLMMs, also known as 'animal models'; Kruuk 2004; Charmantier et al. 2014). These models partition variance in observed phenotypes and, given an appropriate relatedness matrix and model, minimize biases in estimates of V_A and r_A stemming from selection (i.e. non-random variation in fitness, Henderson 1973; Kruuk 2004; Hadfield 2008). Such QGGLMMs can also directly estimate mean additive genetic values of immigrants relative to natives and estimate inbreeding depression, thereby elucidating key roles of immigration and resulting gene flow in shaping latent-scale and phenotypic means and variances (Reid and Keller 2010; Wolak and Keller 2014; Wolak and Reid 2017).

However, despite the recognized need and available statistical methods, few studies have rigorously estimated sex-specific V_As and the cross-sex r_A in fitness in wild populations (Burt 1995; Gardner et al. 2005; Kruuk et al. 2008; Kirkpatrick 2009; Shaw and Shaw 2014; Hendry et al. 2018). Of 17 known studies that estimated V_A for sex-specific absolute fitness measured approximately from zygote to zygote, only eight considered male fitness alongside female fitness (Appendix S1). Since most such studies estimated at least one sex-specific V_A to be close to zero, only two attempted to estimate the cross-sex r_A (McFarlane et al. 2014; Zietsch et al. 2014,

Appendix S1). Two further studies attempted to estimate the cross-sex r_A for fitness measured as an adult's number of adult (i.e. recruited) offspring (Brommer et al. 2007; Foerster et al. 2007, Appendix S1). However, these cross-generation measures of fitness are harder to reconcile with evolutionary theory and risk confounding within-generation selection with evolution (Arnold and Wade 1984; Wolf and Wade 2001). Overall, existing estimates of the cross-sex r_A are very imprecise (Appendix S1). Further, few studies explicitly estimated V_A on appropriate latent scales (but see Milot et al. 2011; McFarlane et al. 2014) and these studies did not attempt to transform latent scale estimates back onto observed phenotypic scales. Such back-transformation is desirable to facilitate cross-study comparison, since latent scale estimates are highly model specific and do not necessarily have a linear relationship to the scale on which phenotypes are expressed and experience natural selection (de Villemereuil et al. 2016). Finally, no studies have yet explicitly estimated additive genetic effects of immigrants, or thereby directly assessed the role of introgressive gene flow in changing local mean breeding value and maintaining evolutionary potential.

The paucity of estimates for sex-specific V_As and cross-sex r_As in fitness likely reflects the substantial challenges of collecting comprehensive sex-specific fitness and relatedness data from free-living individuals. Since all conceived zygotes can rarely be counted, fitness can be pragmatically quantified as the total number of offspring produced over an individual's lifetime, where focal individuals and their offspring are censused as close to conception as feasible (typically soon after birth, hatch, or seed formation, Appendix S2.1). However, most field datasets have some degree of missing or incorrect parentage assignment, and resulting pedigree error could bias quantitative genetic analyses (Brommer et al. 2007; Firth et al. 2015; Wolak and Reid 2017). Further, challenges of tracking juveniles and of paternity assignment mean that records of survival to maturity and male reproductive success are often missing or incorrect (Kruuk et al. 2000; Brommer et al. 2007; Stinchcombe 2014). Observed fitness distributions may also exclude non-

breeders, and hence inaccurately reflect frequencies of individuals with zero fitness (Lebigre et al. 2012). Such error will likely bias key parameter estimates for fitness (e.g. V_A, phenotypic means and variances) and hence bias standardized metrics that depend on such parameters (heritability, h^2; evolvability, I_A; coefficient of additive genetic variance, CV_A; e.g. Freeman-Gallant et al. 2005). Even given comprehensive data spanning multiple generations, V_As and cross-sex r_As in non-Gaussian traits are notoriously difficult to estimate precisely (Shaw 1987; Poissant et al. 2010; Kruuk et al. 2014). Statistical methods that adequately quantify uncertainty should then be used to draw appropriate inference, and thereby facilitate interpretation and subsequent meta-analyses (Garcia-Gonzalez et al. 2012).

To achieve these aims, we fitted Bayesian QGGLMMs to comprehensive multi-generation fitness and pedigree data from song sparrows (*Melospiza melodia*) to estimate sex-specific V_As and r_As (and associated uncertainty) in fitness, and in two hierarchical levels of fitness components. First, we estimated V_A in sex-specific fitness and the cross-sex r_A, thereby evaluating the overall potential for evolutionary change and associated scope for inter-sexual conflict. Second, we estimated V_A and r_A in and among juvenile survival and sex-specific adult LRS, comprising the primary fitness components that generate the distribution of overall fitness. Third, we estimated sex-specific V_A and the cross-sex r_A in adult annual reproductive success (ARS) and estimated V_A in adult annual survival, representing the key fitness components that generate adult LRS. In all cases, we explicitly estimated additive genetic effects of immigrants relative to defined local population founders and estimated inbreeding depression, and thereby evaluated concurrent impacts of natural immigration and resulting gene flow on local additive genetic and phenotypic variation in fitness.

Materials and Methods

STUDY SYSTEM

Estimating V_A and r_A in fitness and fitness components in the wild is perhaps most tractable in populations with limited emigration but sufficient immigration to generate substantial variance in relatedness, and where all local residents and immigrants can be observed. A population of song sparrows inhabiting Mandarte Island, British Columbia, Canada, fulfills these criteria and has proved valuable for quantifying fitness of residents and immigrants and for pedigree-based quantitative genetic analyses (Keller 1998; Marr et al. 2002; Reid et al. 2011a, 2014a,b; Reid and Sardell 2012; Wolak and Reid 2016).

Mandarte's song sparrows typically form socially monogamous breeding pairs, starting from age one year, with a mean of 28±11SD (range 11-52) breeding females per year during 1993-2015. Pairs can rear up to three broods of chicks per year (mean brood size 2.8±1.0SD chicks, range 1-4). However, 28% of offspring are sired by extra-pair males (Sardell et al. 2010), creating opportunities for individual males to gain or lose substantial reproductive success compared to their socially-paired female (Reid et al. 2011a, 2014b; Reid and Sardell 2012). Further, since the adult sex-ratio is often male-biased (mean proportion males during 1993-2015: 0.60±0.09SD, range 0.39-0.75), some males remain socially unpaired in some years (Lebigre et al. 2012), and these males typically gain little extra-pair paternity (Sardell et al. 2010). Consequently, the population's mating system and ecology fosters different means and variances in female versus male reproductive success (Lebigre et al. 2012), creating potential for sexual conflict and trade-offs over fitness components despite social monogamy.

Since 1975, virtually all song sparrow breeding attempts on Mandarte have been closely monitored and all chicks surviving to ca. 6 days post-hatch were marked with unique combinations of metal and colored plastic bands (Smith et al. 2006). Mandarte lies within a large song sparrow meta-population and receives occasional immigrants (totaling 28 females and 16 males during 1976-2014) that were mist-netted and color-banded soon after arriving (Marr et al. 2002; Reid et al.

2006; Smith et al. 2006). Consequently, every song sparrow in the population is individually identifiable by field observation. Comprehensive surveys undertaken each April identified all surviving individuals, including unpaired males, with resighting probability >0.99 (Wilson et al. 2007). Local chick survival from banding to adulthood the following April, and adult survival to subsequent years, were consequently accurately recorded (Keller 1998; Smith et al. 2006).

Each year, the socially-paired parents that reared all banded offspring were identified. To determine genetic parentage, since 1993 all banded chicks and adults were blood sampled and genotyped at 160 polymorphic microsatellite loci. All chicks were assigned to genetic parents with >99% individual-level confidence (Sardell et al. 2010; Nietlisbach et al. 2015). These analyses demonstrated zero extra-pair maternity, and effectively eliminated paternity error. Each banded individual's sex was determined from adult reproductive behavior and/or by genotyping the chromobox-helicase-DNA-binding (CHD) gene (Postma et al. 2011; Nietlisbach et al. 2015).

The local fitness of each chick banded on Mandarte since 1993 was measured as its total lifetime number of chicks banded on Mandarte. Focal chicks that died before adulthood were assigned a fitness of zero (Appendix S2). The two major fitness components, juvenile survival and adult LRS, were respectively measured as survival from banding to adulthood the following April, and as the total number of banded chicks assigned to individuals that survived to adulthood. For each adult, LRS was then further subdivided into ARS and annual survival, respectively measured as the number of banded chicks assigned to each individual in any one year, and as survival to the following April. Since adult (breeding) dispersal away from Mandarte is probably very rare, observed local adult survival likely equates to true survival (Marr et al. 2002; Smith et al. 2006). The relatively high local recruitment rate implies that juvenile (natal) dispersal is also relatively infrequent, although probably non-zero. However, surveys of immediately surrounding islands have detected few local dispersers, implying that unobserved dispersal from Mandarte is likely to be

longer distance. Observed juvenile survival on Mandarte is therefore an appropriate measure of effective local survival and hence local fitness.

QUANTITATIVE GENETIC MODELS

We built a hierarchy of three sets of QGGLMMs designed to estimate sex-specific additive genetic variances (V_A) and covariances (COV_A), and associated standardized statistics (r_A, h^2, I_A, CV_A), in (*i*) sex-specific fitness, then (*ii*) the two major multiplicative components of fitness, namely juvenile survival and adult LRS, then (*iii*) the two components of adult LRS, namely adult ARS and annual survival. Since all traits showed non-Gaussian distributions, all QGGLMMs estimated parameters on latent scales and used appropriate error distributions.

First, we fitted a bivariate QGGLMM to estimate V_A in female and male fitness and the cross-sex COV_A, assuming Poisson distributions with log link functions. Random hatch-year effects were fitted to estimate sex-specific cohort variances in fitness and the cross-sex cohort covariance. Sex-specific residual variances, which measure random individual phenotypic deviations from expectation, were estimated assuming additive overdispersion. Residual covariance was fixed to zero, because there can be no cross-sex covariance between individual phenotypic deviations in traits with sex-limited phenotypic expression. This QGGLMM structure adequately models the observed distributions of fitness while facilitating direct biological interpretation of key parameters. Alternative modeling frameworks with more complex error distributions cannot currently be fitted to wild population relatedness structures and/or do not estimate parameters that directly relate to evolutionary quantitative genetic theory (Appendix S2.2).

Second, we fitted a trivariate QGGLMM to estimate V_A in juvenile survival and adult female and male LRS, and the three pairwise COV_As, thereby considering the major components of overall fitness. We modeled juvenile survival as a single joint trait of both sexes with sex-specific

intercepts, rather than as two sex-specific traits. This simplification facilitated multivariate analysis of juvenile survival alongside sex-specific adult LRS, and is acceptable because previously published and exploratory analyses demonstrated a positive cross-sex r_A for juvenile survival and similar magnitudes of V_A in both sexes, implying moderate shared V_A (Reid and Sardell 2012, Appendix S7). Under these conditions, modeling a single trait for both sexes does not bias estimates of V_A (Wolak et al. 2015), given the degree of uncertainty with which all parameters are estimated. Juvenile survival was modeled as a binary trait with logit link function and residual variance fixed to one. We assumed Poisson distributions for female and male LRS, with log link functions and independent residual variances (as for fitness). Random hatch-year effects were again fitted, thereby estimating cohort variances and covariances in and among the three traits.

Third, we fitted two separate QGGLMMs to estimate V_A in the two major components that generate adult LRS, namely adult ARS and annual survival. For ARS, we fitted a bivariate QGGLMM that estimated V_A in female and male ARS and the cross-sex COV_A, again assuming Poisson distributions for both traits, log link functions, and independent residual variances. Random individual effects were fitted to estimate sex-specific permanent individual variances (i.e. repeatable among-individual variation stemming from environmental and/or non-additive genetic effects). Random year of observation effects were also fitted to estimate among-year environmental variances and the cross-sex year covariance.

For survival, we fitted a univariate QGGLMM that estimated V_A in adult annual survival modeled as a single trait for both sexes with sex-specific intercepts (as for juvenile survival, Appendix S4). We modeled survival as a binary trait expressed by each individual adult in each year, with logit link function and residual variance fixed to one (e.g. Hadfield et al. 2013). Random year of observation and individual effects were fitted to estimate among-year environmental variance and account for overdispersion compared to the assumed geometric distribution of age-specific survival events. Initial models estimated little V_A in adult annual survival, implying that

there can be no genetic covariance (or trade-off) with ARS. Exploratory trivariate QGGLMMs confirmed this view (Appendix S9). Hence, for simplicity, we present separate models for ARS and adult annual survival.

IMMIGRANTS, INBREEDING DEPRESSION, AND FIXED EFFECTS

Standard QGGLMMs estimate V_A and COV_A for a default base population that comprises 'phantom parents' of all pedigreed individuals with unknown parents (Kruuk 2004; Wolak and Reid 2017). In populations with complete local pedigree data for a focal study period but that are open to immigration, the default base population comprises phantom parents of all adults alive at the study start (hereafter 'founders') and of subsequent immigrants. To directly estimate the difference in mean additive genetic value for fitness and fitness components between the defined founders and subsequent immigrants, and account for heterogeneity that could otherwise bias V_A estimates, all QGGLMMs included trait-specific linear regressions on individual immigrant genetic group (*IGG*) coefficient. Each individual's *IGG* coefficient quantifies the expected proportion of that individual's autosomal genome that originated from the defined immigrant group, calculated from pedigree data (Appendix S3). The regression slope (β_{IGG}), modeled as a fixed effect, estimates the difference in mean additive genetic value of the immigrant group relative to the founder group (Wolak and Reid 2017). Since immigration was infrequent, phantom parents of female and male immigrants that arrived in all years were pooled into a single genetic group (Appendix S3). This assumes that the phantom mothers of female and male immigrants have similar mean genetic values as the phantom fathers for any focal trait, and hence that alleles originating in immigrants of both sexes similarly affect the genetic values of descendants of both sexes. This mirrors the standard QGGLMM assumption that phantom mothers and fathers of founders have the same mean breeding values for any focal trait (Wolak et al. 2015).

To quantify inbreeding depression, and minimize bias in V_A estimates that can result from correlated inbreeding across relatives, all four QGGLMMs also included trait-specific linear regressions on individual coefficient of inbreeding (*f*), calculated from pedigree data (Reid and Keller 2010; Wolak and Keller 2014). Regression slopes (β_f) equate to haploid inbreeding load for traits modeled with log link functions, but not with logit link functions (Nietlisbach et al. 2018).

Further fixed effects were restricted to those required to standardize trait observations across individuals. Since juvenile survival probability decreases with increasing seasonal hatch date (Smith et al. 2006), and hatch date reflects the parents' breeding phenotype, models for juvenile survival included a linear regression on the first egg lay date in the nest in which each focal individual hatched. Since adult ARS and annual survival vary with age (Smith et al. 2006; Keller et al. 2008), associated models included categorical effects of age at observation (ages 1, 2, 3-5, or ≥6 years).

PEDIGREE DATA AND MODEL IMPLEMENTATION

Comprehensive pedigree data were initially compiled by assigning all offspring banded during 1975-2014 to their observed socially-paired parents. Paternal links for all chicks hatched during 1993-2014, and 37 additional chicks hatched during 1991-1992, were then corrected for extra-pair paternity based on genotypes at 160 microsatellite loci (Sardell et al. 2010; Reid et al. 2011a; Nietlisbach et al. 2015, 2017). For each QGGLMM, the pedigree was pruned to individuals with observed phenotypes and their known ancestors. The inverse numerator relatedness matrix, and individuals' *IGG* and *f* coefficients, were computed using standard algorithms (Wolak and Reid 2017, Appendix S3). Immigrants were defined as unrelated to all Mandarte residents at arrival, and to subsequent immigrants (Marr et al. 2002; Reid et al. 2006).

For each model, phenotypic data were restricted to cohorts for which all or virtually all individuals had complete fitness or fitness component data, known sex, and genetically verified parents (Appendix S2). Observations of immigrants' own phenotypes were excluded because they

might reflect ecological effects associated with dispersal or subsequent settlement (Marr et al. 2002), and because immigrants' pedigree *f* values are undefined relative to the Mandarte pedigree base population (Reid et al. 2006). However, immigrants that produced ≥1 banded offspring were explicitly included in the pedigree to enable estimation of relatedness among descendants and genetic group effects.

To facilitate estimation for non-Gaussian traits, and associated uncertainty, all models were implemented in a Bayesian framework, using a Markov chain Monte Carlo (MCMC) algorithm to sample posterior distributions. We used diffuse normal prior distributions for all fixed effects (mean=0, variance=10^{10}), and multivariate parameter expanded priors for covariance matrices that gave uniform marginal prior distributions on the correlation. Parameter expanded priors were used for other variance components, giving scaled non-central F-distributions with numerator and denominator degrees of freedom of one (Gelman 2006; Hadfield 2010) and scale parameter of 10 for binary traits or 1,000 for Poisson traits (Appendix S4).

We retained 5,000 samples of each marginal posterior distribution, with MCMC burn-in and thinning interval set to yield absolute autocorrelation values <0.1 and satisfy convergence criteria (Appendix S4). Inference from such posterior distributions should be drawn from defined summary statistics, not directly from the full sample distribution (King et al. 2009 p.85). Posterior distributions can show skew, kurtosis or multiple peaks, including when parameters are near their boundary (e.g. variance near zero). Inferences drawn from posterior modes versus means may then differ. Consequently, we report the marginal posterior mean, mode, and 95% highest posterior density credible intervals (95%CI) and, for key metrics, also depict full marginal posterior distributions alongside prior distributions to further facilitate interpretation (Appendix S4). The 95%CI is especially pertinent when posterior distributions are non-Gaussian and/or uncertainty is large, and directly identifies the parameter values that can and cannot be excluded with 95% confidence given the data, prior, and model (King et al. 2009 pp.86-88).

All QGGLMMs assumed Poisson or binary distributions and therefore estimated (co)variances on latent scales. Posterior distributions of latent-scale heritability (h^2_{latent}) and r_A were computed from all samples of the marginal posterior distributions of underlying components following standard formulae (Appendix S4). Further, to facilitate future comparative studies and evolutionary inferences, we attempted to back-transform posterior distributions of latent-scale variances to the observed phenotypic scale and calculate observed-scale posterior distributions of standardized summary statistics ($h^2_{observed}$, $I_{A\text{-}observed}$, $CV_{A\text{-}observed}$, Appendices S4, S5). However, we could not recover reliable observed-scale variance component posteriors from our bivariate QGGLMM of female and male fitness due to the substantial overdispersion (Appendix S2). $I_{A\text{-}observed}$ was not calculated for juvenile or adult survival because mean standardized variances are not meaningful for binary traits where the mean phenotype is bounded by 0 and 1 (Houle 1992).

Analyses were conducted in R (v3.2.3, R Core Team 2015) using the MCMCglmm (v2.22.1, Hadfield 2010), nadiv (v2.14.3.2, Wolak 2012) and QGglmm (v0.6.0, de Villemereuil et al. 2016) packages. Additional univariate QGGLMMs for sex-specific fitness, and univariate and bivariate QGGLMMs for combinations of juvenile survival and adult LRS, and trivariate models for adult ARS and annual survival (Appendix S9), gave quantitatively similar variance component estimates as the main QGGLMMs presented. Key (co)variance component estimates are robust to reasonable alternative priors (Appendix S6), and remained similar when additional parental and common environmental effects were modeled (Appendix S5). Additional details of model specifications, results, and descriptive figures, are in Appendices S4 and S5. Data and R code for all analyses are available from GitHub: https://github.com/matthewwolak/Wolak_etal_SongSparrowFitnessQG and the Dryad Digital Repository: https://doi.org/10.5061/dryad.p7p1jb3 (Wolak et al. 2018).

Results

348 **FITNESS**

Across 1406 female and 1415 male chicks banded on Mandarte during 1993-2012, 1177 (83.7%)
350 and 1185 (83.7%) respectively had zero fitness. Consequently, fitness distributions were strongly right-skewed, with maxima of 50 and 69 banded offspring for females and males respectively (Fig.
352 1A). Raw mean sex-specific fitness was 1.78 and 1.70 respectively, with substantial phenotypic variances (females 29.8, males 31.7).

354 In the bivariate QGGLMM, the posterior distributions for latent-scale V_A in female and male fitness showed clear peaks that were substantially shifted away from zero and from the prior
356 distributions, indicating substantial V_A for sex-specific fitness (Figs. 2A,B). The posterior modes were similar in both sexes, and the lower 95%CI limits did not converge towards zero (Table 1).
358 There was non-zero cohort variance and substantial residual variance in both sexes, reflecting the overdispersed phenotypic distributions (Table 1, Fig. 1). Consequently, there was relatively small
360 but non-zero heritability of fitness in both sexes; posterior modes and means for h^2_{latent} were 0.08-0.09, with lower 95%CI limits that did not converge to zero (Table 1, Fig. S2).

362 The posterior mode for the cross-sex COV_A in fitness was positive, generating a posterior mode for the cross-sex r_A of intermediate magnitude between zero and one (Table 1, Fig. 2C). The
364 95%CI for r_A was wide and included zero. However, 88% of the posterior density exceeded zero, representing substantial divergence from the uniform prior density, yet the upper 95%CI limit did
366 not converge towards one (Table 1, Fig. 2C). This implies that fitness variation most probably has some, but not all, of the same additive genetic basis in females and males.

368 In total, 26 immigrants that arrived on Mandarte during 1976-2012 made non-zero expected genetic contribution to the 2821 Mandarte-hatched individuals whose fitness was observed
370 (Appendix S3). Across these 2821 individuals, mean *IGG* coefficient was 0.52±0.13SD (range 0.14-0.86). Approximately half the focal individuals' genomes are therefore expected to have
372 originated from immigrants on average, implying that immigration could contribute substantially to

standing V_A within the Mandarte breeding population. The posterior modes for the regressions of sex-specific fitness on *IGG*, which quantify mean immigrant genetic group effects, were negative in both sexes with 95%CIs that did not overlap zero (Table 1). Additive effects of alleles carried by immigrants therefore decreased fitness, relative to additive effects of alleles in the defined founder population, in both sexes.

Across the 2821 individuals, mean *f* was 0.074±0.052 (range 0.000-0.347, 7.4% zeroes). Substantial variation in *f* was directly attributable to immigration: 91% of individuals with *f*=0 had one immigrant parent. However, since immigrants' descendants commonly inbred in future generations, the model covariates *f* and *IGG* were only moderately correlated across individuals (Pearson correlation coefficients: females r=-0.25, males r=-0.30). The posterior modes for the regressions of sex-specific fitness on *f* were negative with 95%CIs that did not overlap zero, demonstrating very strong inbreeding depression in fitness in both sexes (Table 1).

JUVENILE SURVIVAL AND ADULT LIFETIME REPRODUCTIVE SUCCESS

Of 1542 female and 1562 male chicks banded during 1993-2014, 254 (16.5%) females and 331 (21.2%) males survived on Mandarte to the following April. Adult LRS was measured for 243 adult females and 312 adult males hatched during 1993-2012, with sex-specific means of 10.3 (median 7, variance 85.1, 5.8% zeroes) and 7.7 (median 4, variance 97.6, 26.3% zeroes) banded offspring respectively (Fig. 1B).

In the trivariate QGGLMM, the posterior distribution for V_A in juvenile survival showed a clear peak, and hence posterior mean, that departed from zero and from the prior distribution. However, the lower 95%CI limit converged towards zero, and there was a second peak of posterior density near zero that mirrored the prior distribution (Table 2, Fig. 3A). Since there was substantial cohort variance (Table 2), the posterior means for h^2_{latent} and $h^2_{observed}$ were small, but again showed clear peaks away from zero (Fig. S3). Although the lower 95%CI limits converged towards zero,

approximately 93% and 82% of posterior samples for h^2_{latent} and $h^2_{observed}$, respectively, exceeded a minimal value of 0.01 (Table 2, Fig. S3).

The posterior mode for V_A in adult female LRS was very small (Table 2). The posterior mean was slightly greater due to the right-skewed posterior distribution (Table 2, Fig. 3B). However there was substantial posterior density close to zero compared to the prior distribution, and the lower 95%CI limit converged towards zero (Fig. 3B, Table 2). Consequently, the posterior modes (and means) of h^2_{latent}, $h^2_{observed}$ and $I_{A\text{-}observed}$ for female LRS were small, with lower 95%CI limits that converged towards zero (Table 2, Figs. S4, S5).

In marked contrast, the posterior mode and mean for V_A in adult male LRS were substantial and the lower 95%CI limit considerably exceeded zero (Table 2, Fig. 3C). Consequently, although there were also moderate cohort and residual variances, the posterior mode and mean for h^2_{latent} for male LRS were substantial (Table 2, Fig. S4). These values were smaller for $h^2_{observed}$, reflecting the non-linear transformation induced by the mean-variance relationship of the Poisson distribution, but the lower 95%CI limit still did not converge towards zero (Table 2, Fig. S4). The posterior mode for $I_{A\text{-}observed}$ for male LRS was also moderate (Table 2, Fig. S5). Overall, there was substantially more V_A in adult male LRS than adult female LRS, as the posterior distribution of the male-female difference in V_A had a posterior mean of 1.14 (mode=0.97, 95%CI limits:0.19, 2.15).

Since V_A in female LRS was so small and the lower 95%CI limit for V_A in juvenile survival also converged towards zero, the pairwise COV_As and r_As among juvenile survival and female and male LRS were unsurprisingly estimated with considerable uncertainty (Table 2, Fig. 3). The posterior modes and means for r_A between juvenile survival and male LRS, and between female and male LRS, were slightly negative, but spanned zero for juvenile survival and female LRS, all with 95%CI limits that did not converge towards either -1 or 1 (Table 2, Fig. 3).

Distributions of *IGG* and *f* for individuals included in analyses of juvenile survival and adult LRS (and ARS and survival) were quantitatively similar to those for individuals included in

analyses of fitness (summarized above). The posterior mode for the regression of juvenile survival on *IGG* was negative, with a 95%CI that did not overlap zero (Table 2). Further analyses showed similar negative slopes for female and male juvenile survival modeled as separate traits (Appendix S7). However, the posterior modes for the regressions of adult female and male LRS on *IGG* were small, with 95%CIs that spanned zero (Table 2). This implies that additive effects of immigrants' alleles decreased local juvenile survival, but not adult female or male LRS, relative to additive effects of founders' alleles.

The posterior modes for the regressions of juvenile survival and adult female and male LRS on *f* were all negative, demonstrating inbreeding depression (although the 95%CI for female LRS overlapped zero, Table 2). Further, inbreeding depression in LRS is most likely stronger in males than females (Table 2), as the posterior distribution of the male-female difference in *f* had a posterior mean of -7.12 (mode=-5.51, 95%CI limits:-14.0, -0.45).

ADULT ANNUAL REPRODUCTIVE SUCCESS

During 1994-2015, there were 526 and 773 observations of ARS for adult females and males respectively, involving 254 and 331 Mandarte-hatched individuals. Mean female ARS was 4.9 banded offspring (median 5, variance 6.2, range 0-11, 6.7% zeroes, Fig. 1C) and mean male ARS was 3.2 banded offspring (median 2, variance 13.2, range 0-21, 32.7% zeroes, Fig. 1C).

In the bivariate QGGLMM, the posterior mode for V_A in female ARS was very small and the lower 95%CI limit converged towards zero (Table 3, Fig. 4A). However, there was a clear secondary peak away from zero (Fig. 4A), meaning that the posterior mean was slightly larger (Table 3), and 75% of the posterior density exceeded a minimal value of 0.01. Further, the posterior density near zero resembles the prior distribution, suggesting that the prior influences the posterior mode, while the data generate the second peak about the posterior mean. This implies the existence of very small, but probably non-zero, V_A for female ARS (Fig. 4A inset).

In contrast, the posterior mode and mean for V_A in male ARS were substantially larger and the lower 95%CI limit did not converge towards zero (Table 3, Fig. 4B). The permanent individual variances were very small in both sexes, but the year and residual variances were substantial, especially for males (Table 3). Consequently, despite the marked difference in V_A, the posterior means for h^2_{latent} and $h^2_{observed}$ for ARS were similar for both sexes (~0.06-0.18), but $I_{A\text{-}observed}$ was substantially greater for male ARS than female ARS (Table 3, Figs. S6, S7).

The posterior mode for the cross-sex additive genetic correlation (r_A) in ARS was positive but small. Due to the small V_A in female ARS, the 95%CI was again wide and spanned zero, but did not converge towards either -1 or 1 (Table 3, Fig. 4C).

The posterior modes for the regressions of ARS on *IGG* were small in both sexes, with 95%CIs that overlapped zero (Table 3). The posterior modes for the regressions of ARS on *f* were negative in both sexes, although the 95%CI for females again overlapped zero (Table 3). Again, inbreeding depression is most likely stronger in male than female ARS (Table 3), as the posterior distribution of the male-female difference in *f* had a posterior mean of -3.69 (mode=-3.43, 95%CI limits:-6.93, -0.57).

ADULT ANNUAL SURVIVAL

For the focal 254 adult females and 331 adult males, the mean number of observations of annual survival (or mortality) was 2.1 (median 1, range 1-9, Fig. 5A) for females and 2.3 (median 2, range 1-9, Fig. 5B) for males, representing overall annual survival of 53.0% and 58.0% respectively.

In the univariate QGGLMM, the posterior mode for V_A was effectively zero (Table 3, Fig. 5C). The posterior mean was slightly larger, but there was substantial posterior density close to zero compared to the prior distribution, and the lower 95%CI limit converged to zero (Table 3). Since there was also substantial year variance, the posterior modes for h^2_{latent} and $h^2_{observed}$ were very small (Table 3; Fig. S8.3). The posterior modes for the regressions of adult annual survival on *IGG* and *f*

were also small, with 95%CIs that overlapped zero (Table 3). Analyses of adult longevity rather than annual survival, and of sex-specific annual survival, yielded similar conclusions (Appendix S8).

Discussion

ADDITIVE GENETIC VARIANCE AND CORRELATION IN SEX-SPECIFIC FITNESS

The sex-specific additive genetic variances (V_A) in fitness, and the cross-sex genetic correlation (r_A), are key parameters that determine the rate of fitness evolution and shape evolutionary responses to natural and sexual selection (Burt 1995; Brommer et al. 2007; Kirkpatrick 2009; Shaw and Shaw 2014). They also underlie the potential for evolutionary sexual conflict, which might constrain evolution yet help maintain overall V_A in fitness (Lande 1980; Chippindale et al. 2001; Kruuk et al. 2008; Bonduriansky and Chenoweth 2009; Long et al. 2012). However, these key parameters have rarely been estimated in wild populations, particularly using theoretically appropriate measures of fitness while accommodating non-Gaussian phenotypic distributions and accounting for genetic effects of immigration and inbreeding (Kruuk et al. 2008; Kirkpatrick 2009; Shaw and Etterson 2012; Gomulkiewicz and Shaw 2013; Shaw and Shaw 2014).

Our analyses of comprehensive fitness data from free-living song sparrows estimated non-zero latent-scale V_As and heritabilities for fitness, of similar magnitudes, in both sexes. Such estimates do not concur with the common assumption that V_A for fitness is usually negligible (Charlesworth 1987; Shaw and Shaw 2014; Walling et al. 2014). Instead, our estimates support the view that non-trivial V_A in fitness can be readily generated and/or maintained in wild populations (e.g. Houle 1992; Kirkpatrick 2009; Zhang 2012; Shaw and Shaw 2014). Further, our inference that the cross-sex r_A for fitness is most likely to be positive implies that some V_A is shared between the sexes, potentially facilitating an increase in population mean fitness (Lande 1980). However, the

498 upper 95%CI limit for the cross-sex r_A in fitness was less than one. For the special case of fitness, this implies that some sex-limited or sexually antagonistic genetic variation does exist, potentially
500 facilitating the maintenance of overall V_A.

The few available estimates of sex-specific V_A in fitness in wild populations cannot readily be
502 compared quantitatively because different studies used different fitness metrics, analytical methods, and estimation scales, with different degrees of paternity error and missing data. However,
504 qualitatively concordant with our results, V_A for fitness was estimated to be non-zero and similar in both sexes in collared flycatchers (*Ficedula albicollis*, Merilä and Sheldon 2000; Brommer et al.
506 2007) and Swedish humans (*Homo sapiens*, Zietsch et al. 2014). In contrast, V_A was estimated to be zero or very small in both sexes in great tits (*Parus major*, McCleery et al. 2004), bighorn sheep
508 (*Ovis canadensis*, Coltman et al. 2005), North American red squirrels (*Tamiasciurus hudsonicus*, McFarlane et al. 2014), and savannah sparrows (*Passerculus sandwichensis*, Wheelwright et al.
510 2014); zero in females but more substantial in males in red deer (*Cervus elaphus*, Kruuk et al. 2000, but see Foerster et al. 2007) and Austrian humans (Gavrus-Ion et al. 2017); yet zero in males but
512 more substantial in females in red-billed gulls (*Larus novaehollandiae*, Teplitsky et al. 2009) and pre-industrial Finnish humans (Pettay et al. 2005, Appendix S1).

514 Meanwhile, the posterior 95%CI for the cross-sex r_A for fitness in song sparrows excluded the substantial negative values previously estimated in wild populations (Foerster et al. 2007; Brommer
516 et al. 2007; McFarlane et al. 2014; Appendix S1), with relatively little posterior density surrounding the small or slightly negative values estimated in laboratory populations (Chippindale et al. 2001;
518 Delcourt et al. 2009; Innocenti and Morrow 2010; Collet et al. 2016). Yet, cross-sex r_As can change substantially when (laboratory) populations experience novel environments (Delcourt et al. 2009;
520 Punzalan et al. 2014; Collet et al. 2016), migration load (Long et al. 2012), or inbreeding (Duffy et al. 2014). Positive values, such as those inferred in the song sparrows, might indicate populations
522 where both sexes are displaced from their fitness peak, and consequently experience congruent

directional selection (Long et al. 2012; Duffy et al. 2014; Punzalan et al. 2014). Overall, further rigorous and standardized estimates of V_A and r_A in sex-specific fitness from wild populations experiencing different ecological circumstances are clearly required to discern general patterns and evolutionary implications.

ADDITIVE GENETIC VARIANCES AND CORRELATIONS IN FITNESS COMPONENTS

Values of V_A in sex-specific fitness, and the cross-sex r_A, must ultimately result from V_As and cross-sex and within-sex r_As in underlying sex-specific fitness components. Quantifying such parameters can consequently help identify mechanisms that maintain V_A in fitness, and identify sources of sexual conflict (Walling et al. 2014). Juvenile survival constitutes one primary fitness component. Indeed, 96% of observed song sparrow fitness values of zero represent individuals that did not (locally) survive to adulthood, generating the high frequency of zeroes in the overall fitness distribution. Such patterns are likely commonplace (Blomquist 2010; Wagenius et al. 2010; Gomulkiewicz and Shaw 2013). We therefore explicitly estimated V_A in juvenile survival, showing a moderate posterior mean with a clear peak of posterior density away from zero, but a lower 95%CI that converged towards zero. This broadly concurs with previous evidence that V_A in juvenile survival is moderate and similar in female and male song sparrows with a substantial positive cross-sex r_A (Reid and Sardell 2012, Appendix S7).

However, for adult LRS, which constitutes the remaining primary fitness component, there was a striking difference between the sexes: V_A for male LRS was substantial and clearly exceeded zero, while V_A for female LRS was very small. This implies that there is little opportunity for rapid evolutionary change in adult female LRS, and hence in other traits that are postulated to be genetically correlated with female reproductive success. The potential for evolution in mean absolute male LRS will consequently also be constrained. Yet, the large V_A in absolute male LRS estimated on the QGGLMM latent scale implies that relative male LRS on the observed scale will

548 also exhibit V_A. Consequently, there could be rapid evolutionary change in relative male LRS, and hence in genetically correlated traits. This includes traits that shape mating systems, for example
550 male acquisition of reproductive success through within-pair versus extra-pair paternity (Reid and Wolak 2018).

552 The small V_A in female LRS impedes precise estimation of the cross-sex r_A in LRS, and indeed renders such estimation somewhat redundant (since r_A is undefined if V_A is truly zero in one or both
554 sexes). Nevertheless, the cross-sex r_A posterior mode was close to zero, further suggesting that additive genetic effects on adult LRS are largely independent in females and males. Together, our
556 results imply that the moderate positive cross-sex r_A in fitness is primarily driven by the moderate positive cross-sex r_A in juvenile survival. Hence, cross-sex expression of additive genetic effects on
558 juvenile survival ameliorates potential sexually antagonistic genetic variation in overall fitness resulting from sex-specific expression of adult LRS. These patterns are reminiscent of those
560 observed in *Drosophila melanogaster*, where a positive cross-sex r_A in juvenile survival initially combined with a negative cross-sex r_A in adult reproductive success to generate a weak overall
562 cross-sex r_A for fitness (Chippindale et al. 2001), but where the cross-sex r_A in adult reproductive success was no longer detectably different from zero after further generations of laboratory
564 adaptation (Collet et al. 2016).

Further decomposition of adult LRS in song sparrows revealed little or no V_A in adult annual
566 survival, and identified ARS as the primary source of V_A in male LRS. The substantial difference in V_A in ARS, and hence LRS, between males and females likely reflects the population's ecology and
568 mating system. Due to the typically male-biased adult sex-ratio and frequent extra-pair paternity, males accumulate ARS by securing a territory and a social mate, defending within-pair paternity
570 and accruing extra-pair paternity (Sardell et al. 2010; Lebigre et al. 2012; Reid et al. 2011a, 2014a,b; Losdat et al. 2015). In contrast, females accumulate ARS through their own fecundity.
572 Consequently, while components of ARS such as within-pair paternity can be conceptualized as

'emergent' traits of pairs rather than individuals (Reid et al. 2014a), males and females are likely to
574 differ substantially in the suite of physiological and behavioral traits that generate high ARS, and hence in underlying genetic effects. Previous analyses revealed non-zero V_A in components of
576 annual male extra-pair reproductive success and a positive r_A with within-pair paternity success per brood (Reid et al. 2014b; Losdat et al. 2015), but a negative r_A between net paternity success and
578 juvenile survival (Reid and Sardell 2012). Together, these positive and negative correlations, alongside among-year variation in adult sex-ratio and hence the social context in which male
580 reproductive success is expressed, could help maintain substantial V_A in male ARS (and hence LRS, Reid and Wolak 2018).

582

GENETIC EFFECTS OF IMMIGRATION

584 Immigration, and resulting gene flow, is one primary mechanism that can maintain V_A in fitness and associated evolutionary potential in any population, and also rapidly increase mean fitness by
586 alleviating inbreeding depression. However the overall genetic effects of immigration, and the evolutionary consequences, depend on genetic properties of naturally-occurring immigrants
588 compared to existing natives (Ingvarsson and Whitlock 2000; Tallmon et al. 2004; Edelaar and Bolnick 2012; Carlson et al. 2014). We utilized the multi-generation song sparrow pedigree, that
590 links all Mandarte-hatched individuals to their immigrant and 'founder' ancestors and hence describes expected introgression of immigrants' alleles, to directly estimate the relative mean
592 additive genetic values for local fitness of the defined immigrant and founder genetic groups. Unlike analyses that examine demographic and evolutionary consequences of dispersal by directly
594 comparing observed phenotypes of immigrants (or dispersers) and residents (e.g. Marr et al. 2002; Pasinelli et al. 2004; Nosil et al. 2005; Pärn et al. 2009; Bonte et al. 2012), our analyses do not
596 utilize immigrants' own phenotypes and consequently cannot be confounded by environmental effects of dispersal on those phenotypes. Our analyses showed that immigrant song sparrows carry

alleles that, when expressed in subsequent Mandarte-hatched generations, have negative additive effects on local fitness in both sexes.

Such effects could stem from three main processes. First, there could be divergent selection among song sparrow demes and resulting local adaptation. Immigrants to Mandarte might consequently be locally maladapted and hence have low mean additive genetic value for local fitness, as assumed by classical migration load models. Second, dispersal could be non-random, such that immigrants have low additive genetic value for local fitness even without any local adaptation. Third, low additive genetic value for fitness measured on Mandarte could reflect V_A in dispersal, such that offspring of immigrants are more likely to emigrate and hence have zero local fitness (e.g. Doligez and Pärt 2008). These three processes are not mutually exclusive and are not distinguished by our current analyses. However, the overall effects resulted primarily from immigrants' low additive genetic value for local juvenile survival, and therefore reflects some combination of effects on early-life mortality and/or emigration. To indicate biological effects, the estimated latent-scale effect of β_{IGG} = -2.6 (Table 2) implies a decrease in local juvenile survival probability of approximately 0.04 given an increase in individual *IGG* coefficient of 0.1 spanning the current mean of ~0.5, which is not trivial. In general, such reduced local survival of immigrants' descendants would reduce the effective rate of gene flow below that expected given the observed immigration rate (Garant et al. 2007).

However, our analyses also demonstrate very strong inbreeding depression in fitness in both sexes, resulting from inbreeding depression in juvenile survival and in adult LRS and ARS, particularly in males. Similar patterns of inbreeding depression have previously been documented in the Mandarte population, using different data subsets and methods (Keller 1998; Reid et al. 2014c; Nietlisbach et al. 2017). This inbreeding depression reflects covariance between individual fitness and *f*, where the underlying variance in *f* stems from immigrant-native outcrossing; resulting F1 offspring are defined as outbred and have relatively high fitness (Keller 1998; Marr et al. 2002;

Reid et al. 2006, 2014c; Wolak and Reid 2016). The estimated latent-scale effect size of β_f = -9.3 (Table 2) implies an increase in juvenile survival probability of approximately 0.25 for outbred offspring (*f*=0) compared to inbred offspring with *f*=0.1 (see also Keller 1998; Reid et al. 2014c). This effect could cause rapid initial introgression of immigrants' alleles, and hence increase the short-term effective rate of gene flow (e.g. Ingvarsson and Whitlock 2000; Garant et al. 2007; Hedrick et al. 2014). Indeed, the mean *IGG* coefficient of ~0.5, calculated across the focal 2821 fitness-phenotyped individuals, implies that an average Mandarte-hatched song sparrow inherited half its genome from immigrant ancestors despite the relatively small number of contributing immigrants (n=26) and that only 195 (7%) of the phenotyped individuals were direct F1 offspring of immigrant-native pairings.

However, once immigrants' descendants start to inbreed, as is inevitable for initially high-fitness lineages in small populations (Reid et al. 2006; Bijlsma et al. 2010; Hedrick et al. 2014), increased expression of recessive alleles with detrimental effects on local fitness would occur. This process would exacerbate the decrease in fitness that is expected following recombination in F2 and subsequent generations and resulting outbreeding depression (Marr et al. 2002; Frankham 2016. The combination of heterosis that exacerbates initial introgression and low overall additive genetic value for fitness could potentially generate substantial migration load; almost all population members might be pulled below the fitness peak, substantially decreasing population mean fitness but potentially generating a positive overall cross-sex r_A for fitness and alleviating sexual conflict (Long et al. 2012; Duffy et al. 2014; Punzalan et al. 2014). Such multi-generational dynamics of immigrants' alleles should, in future, be explicitly quantified using pedigree and genomic data (from song sparrows and other systems), and through theory that simultaneously considers heterosis and migration load (e.g. Lopez et al. 2009). Meanwhile, our analyses demonstrate that structured quantitative genetic analyses can explicitly estimate V_A in fitness alongside multiple genetic consequences of immigration in wild populations, and thereby elucidate the contributions of gene

flow to the magnitude and maintenance of overall V_A in fitness and resulting evolutionary dynamics.

LITERATURE CITED

Arnold, S. J., and M. J. Wade. 1984. On the measurement of natural and sexual selection: Applications. Evolution 38:720–734.

Barton, N. H., and P. D. Keightley. 2002. Understanding quantitative genetic variation. Nat. Rev. Gen. 3:11-21.

Bell, G. 2013. Evolutionary rescue and the limits of adaptation. Philos. Trans. R. Soc. B 368:20120080.

Bijlsma, R., M. D. D. Westerhof, L. P. Roekx, and I. Pen. 2010. Dynamics of genetic rescue in inbred Drosophila melanogaster population. Cons. Genet. 11:449-462.

Blomquist, G.E. 2010. Heritability of fitness in female macaques. Evol. Ecol. 24:657-669.

Bonduriansky, R., and S. F. Chenoweth. 2009. Intralocus sexual conflict. Trends Ecol. Evol. 24:280-288.

Bonte, D. et al. 2012. Costs of dispersal. Biol. Rev. 87:290-312.

Brommer, J. E. 2000. The evolution of fitness in life-history theory. Biol. Rev. 75:377–404.

Brommer, J. E., M. Kirkpatrick, A. Qvarnström, and L. Gustafsson. 2007. The intersexual genetic correlation for lifetime fitness in the wild and its implications for sexual selection. PLoS One 2:e744.

Burt, A. 1995. The evolution of fitness. Evolution, 49:1-8.

Carlson, S. M., C. J. Cunningham, and P. A. H. Westley. 2014. Evolutionary rescue in a changing world. Trends. Ecol. Evol. 29:521-530.

672 Charlesworth, B. 1987. The heritability of fitness. Pp. 21–40 *in* J. W. Bradbury and M. B. Andersson, eds. Sexual Selection: Testing the Alternatives. Wiley, Chichester.

674 Charmantier, A., D. Garant, and L. E. B. Kruuk (eds). 2014. Quantitative Genetics in the Wild. Oxford University Press, Oxford.

676 Chippindale, A. K., J. R. Gibson, and W. R. Rice. 2001. Negative genetic correlation for adult fitness between sexes reveals ontogenetic conflict in *Drosophila*. Proc. Natl. Acad. Sci. U. S. A. 98:1671–1675.

678 Collet, J. M., S. Fuentes, J. Hesketh, M. S. Hill, P. Innocenti, E. H. Morrow, K. Fowler, and M. Reuter. 2016. Rapid evolution of the intersexual genetic correlation for fitness in *Drosophila melanogaster*. Evolution 70:781-795.

680 Coltman, D. W., P. O'Donoghue, J. T. Hogg, and M. Festa-Bianchet. 2005. Selection and genetic (co)variance in bighorn sheep. Evolution 59:1372-1382.

682 Crow, J. F., and M. Kimura. 1970. Introduction to Population Genetics Theory. Harper and Row, New York.

684 de Villemereuil, P., H. Schielzeth, S. Nakagawa, and M. Morrissey. 2016. General methods for evolutionary quantitative genetic inference from generalized mixed models. Genetics 204:1281–1294.

686 Delcourt, M., M. W. Blows, and H. D. Rundle. 2009. Sexually antagonistic genetic variance for fitness in an ancestral and a novel environment. Proc. R. Soc. B 276:2009–14.

688 Doligez, B., and T. Pärt. 2008. Estimating fitness consequences of dispersal: a road to 'know-where'? Non-random dispersal and the underestimation of dispersers' fitness. J. Anim. Ecol. 77:1199-1211.

690 Duffy, E., R. Joag, J. Radwan, N. Wedell, and D. J. Hosken. 2014. Inbreeding alters intersexual fitness correlations in *Drosophila simulans*. Ecol. and Evol. 4:3330-3338.

696 Edelaar, P., and D. I. Bolnick. 2012. Non-random gene flow: an underappreciated force in evolution and ecology. Trends Ecol. Evol. 27:659-665.

698 Ellegren, H., and B. C. Sheldon. 2008. Genetic basis of fitness differences in natural populations. Nature 452:169-175.

700 Falconer, D. S. 1989. Introduction to Quantitative Genetics. 3rd ed. John Wiley & Sons, Inc., New York.

702 Firth, J. A., J. D. Hadfield, A. W. Santure, J. Slate, and B. C. Sheldon. 2015. The influences of nonrandom extra-pair paternity on heritability estimates derived from wild pedigrees.

704 Evolution 69:1336-1344.

Fisher, R. A. 1930. The Genetical Theory of Natural Selection. Clarendon Press, Oxford.

706 Flint, J., and T. F. C. Mackay. 2009. Genetic architecture of quantitative traits in mice, flies, and humans. Genome Res. 19:723–733.

708 Foerster, K., T. Coulson, B. C. Sheldon, J. M. Pemberton, T. H. Clutton-Brock, and L. E. B. Kruuk. 2007. Sexually antagonistic genetic variation for fitness in red deer. Nature 447:1107-1111.

710 Frankham, R. 2016. Genetic rescue benefits persist to at least the F3 generation, based on meta-analysis. Biol. Cons. 195, 33-36

712 Freeman-Gallant, C. R., N. T. Wheelwright, K. E. Meiklejohn, S. L. States, and S. V Sollecito. 2005. Little effect of extrapair paternity on the opportunity for sexual selection in Savannah

714 sparrows (*Passerculus sandwichensis*). Evolution 59:422–430.

Garant, D., S. E. Forde, and A. P. Hendry. 2007. The multifarious effects of dispersal and gene flow

716 on contemporary adaptation. Func. Ecol. 21:434-443.

Garcia-Gonzalez, F., L. W. Simmons, J. L. Tompkins, J. S. Kotiaho, and J. P. Evans. 2012.

718 Comparing evolvabilities: common errors surrounding the calculation and use of coefficients of additive genetic variation. Evolution 60, 2341-2349.

Gardner, M. P., K. Fowler, N. H. Barton, and L. Partridge. 2005. Genetic variation for total fitness in *Drosophila melanogaster*: Complex yet replicable patterns. Genetics 169:1553–1571.

Gavrus-Ion, A., T. Sjøvold, M. Hernández, R. González-José, M. E. Esteban Torné, N. Martínez-Abadías, and M. Esparza. 2017. Measuring fitness heritability: Life history traits versus morphological traits in humans. Am. J. Phys. Anthropol. 164:321-330.

Gelman, A. 2006. Prior distributions for variance parameters in hierarchical models. Bayesian Anal. 1:515–533.

Gomulkiewicz, R., and R. G. Shaw. 2013. Evolutionary rescue beyond the models. Philos. Trans. R. Soc. B 368:20120093.

Hadfield, J. D. 2008. Estimating evolutionary parameters when viability selection is operating. Proc. R. Soc. B 275:723–734.

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

Hadfield, J. D., E. A. Heap, F. Bayer, E. A. Mittell, and N. M. A. Crouch. 2013. Disentangling genetic and prenatal sources of familial resemblance across ontogeny in a wild passerine. Evolution 67:2701–2713.

Hedrick, P. W., R. O. Peterson, L. M. Vucetich, J. R. Adams, and J. A. Vucetich. 2014. Genetic rescue in Isle Royale wolves: genetic analysis and the collapse of the population. Cons. Genet. 15:1111-1121.

Henderson, C. R. 1973. Sire evaluation and genetic trends. J. Anim. Sci. 1973:10–41.

Hendry, A. P., D. J. Schoen, M. E. Wolak, and J. M. Reid. 2018. The contemporary evolution of fitness. Ann. Rev. Ecol. Evol. Syst. *In press*. https://doi.org/10.1146/annurev-ecolsys-110617-062358.

Hill, W. G. 2012. Quantitative genetics in the genomics era. Curr. Genomics 13:196–206.

744 Houle, D. 1992. Comparing evolvability and variability of quantitative traits. Genetics 130:195–
204.
746 Innocenti, P., and E. H. Morrow. 2010. The sexually antagonistic genes of *Drosophila*
melanogaster. PLoS Biol. 8:e1000335.
748 Ingvarsson, P. K., and M. C. Whitlock. 2000. Heterosis increases the effective migration rate. Proc.
R. Soc. B. 267:1321-1326.
750 Johnston, S. E., J. Gratten, C. Berenos, J. G. Pilkington, T. H. Clutton-brock, J. M. Pemberton, and
J. Slate. 2013. Life history trade-offs at a single locus maintain sexually selected genetic
752 variation. Nature 502:93–95.
Keller, L. F. 1998. Inbreeding and its fitness effects in an insular population of song sparrows
754 (*Melospiza melodia*). Evolution 52:240–250.
Keller, L. F., J. M. Reid, and P. Arcese. 2008. Testing evolutionary models of senescence in a
756 natural population: age and inbreeding effects on fitness components in song sparrows. Proc.
R. Soc. B 275:597-604.
758 King, R., B. J. T. Morgan, O. Gimenez, and S. P Brooks. 2009. Bayesian Analysis for Population
Ecology. Chapman & Hall.
760 Kirkpatrick, M. 2009. Patterns of quantitative genetic variation in multiple dimensions. Genetica
136:271-284.
762 Kruuk, L. E. B., T. H. Clutton-Brock, J. Slate, J. M. Pemberton, S. Brotherstone, and F. E.
Guinness. 2000. Heritability of fitness in a wild mammal population. Proc. Natl. Acad. Sci.
764 97:698-703.
Kruuk, L. E. B., A. Charmantier, and D. Garant. 2014. The study of quantitative genetics in wild
766 populations. Pp. 1–15 in A. Charmantier, D. Garant, and L. E. B. Kruuk, eds. Quantitative
Genetics in the Wild. Oxford University Press, Oxford.
768

Kruuk, L. E. B. 2004. Estimating genetic parameters in natural populations using the "animal model." Philos. Trans. R. Soc. B 359:873–890.

Kruuk, L. E. B., J. Slate, and A. J. Wilson. 2008. New answers for old questions: The evolutionary quantitative genetics of wild animal populations. Annu. Rev. Ecol. Evol. Syst. 39:525–548.

Lande, R. 1980. Sexual dimorphism, sexual selection, and adaptation in polygenic characters. Evolution 34:292–305.

Lande, R. 1982. A quantitative genetic theory of life history evolution. Ecology 63:607–615.

Lebigre, C., P. Arcese, R. J. Sardell, L. F. Keller, and J. M. Reid. 2012. Extra-pair paternity and the variance in male fitness in song sparrows (*Melospiza melodia*). Evolution 66:3111–3129.

Lenormand, T. 2002. Gene flow and the limits to natural selection. Trends Ecol. Evol. 17:183-189.

Lewontin, R. C. 1974. Genetic Basis of Evolutionary Change. Columbia University Press, New York.

Long, T. A. F., A. F. Agrawal, and L. Rowe. 2012. The effect of sexual selection on offspring fitness depends on the nature of genetic variation. Cur. Biol. 22:204-208.

Lopez, S., F. Rousset, F. H. Shaw, R. G. Shaw, and O. Ronce. 2009. Joint effects of inbreeding and local adaptation on the evolution of genetic load after fragmentation. Cons. Biol. 23:1618-1627.

Losdat, S., P. Arcese, and J. M. Reid. 2015. Double decomposition: decomposing the variance in subcomponents of male extra-pair reproductive success. J. Anim. Ecol. 84:1384-1395.

Lynch, M., and B. Walsh. 1998. Genetics and analysis of quantitative traits. Sinauer, Sunderland, USA.

Marr, A. B., L. F. Keller, and P. Arcese. 2002. Heterosis and outbreeding depression in descendants of natural immigrants to an inbred population of song sparrows (Melospiza melodia). Evolution 56:131-142.

McCleery, R. H., R. A. Pettifor, P. Armbruster, K. Meyer, B. C. Sheldon, and C. M. Perrins. 2004. Components of variance underlying fitness in a natural population of the great tit *Parus major*. Am. Nat. 164:E62-E72.

McFarlane, S. E., J. C. Gorrell, D. W. Coltman, M. M. Humphries, S. Boutin, and A. G. McAdam. 2014. Very low levels of direct additive genetic variance in fitness and fitness components in a red squirrel population. Ecol. & Evol. 4:1729-1738

Merilä, J., and B. C. Sheldon. 1999. Genetic architecture of fitness and nonfitness traits: empirical patterns and development of ideas. Heredity 83:103–109.

Merilä, J., and B. C. Sheldon. 2000. Lifetime reproductive success and heritability in nature. Am. Nat. 155:301-310.

Metcalf, C. J. E., and S. Pavard. 2007. Why evolutionary biologists should be demographers. Trends Ecol. Evol. 22:205-212.

Milot, E., F. M. Mayer, D. H. Nussey, M. Boisvert, F. Pelletier, and D. Réale. 2011. Evidence for evolution in response to natural selection in a contemporary human population. Proc. Natl. Acad. Sci. 108:17040–45

Nietlisbach, P., G. Camenisch, T. Bucher, J. Slate, L. F. Keller, and E. Postma. 2015. A microsatellite-based linkage map for song sparrows (*Melospiza melodia*). Mol. Ecol. Res. 15:1486–1496.

Nietlisbach, P., L. F. Keller, G. Camenisch, F. Guillaume, P. Arcese, J. M. Reid, and E. Postma. 2017. Pedigree-based inbreeding coefficient explains more variation in fitness than heterozygosity at 160 microsatellites in a wild bird population. Proc. R. Soc. B 284:20162763.

Nietlisbach, P., S. Muff, J. M. Reid, M. C. Whitlock, and L. F. Keller. *In review*. Nonequivalent lethal equivalents: statistical approaches for comparable estimation of inbreeding depression with pedigree-based and genomic metrics of inbreeding. Evol. App.

Nosil, P., T. H. Vines, and D. J. Funk. 2005. Perspective: Reproductive isolation caused by natural selection against immigrants from divergent habitats. Evolution 59:705-719.

Orr, H. A. 2009. Fitness and its role in evolutionary genetics. Nat. Rev. Genet. 10:531–539.

Pärn, H., H. Jensen, T. H. Ringsby, and B.-E. Sæther. 2009. Sex-specific fitness correlates of dispersal in a house sparrow metapopulation. J. Anim. Ecol. 78:1216-1225.

Pasinelli, G., K. Schiegg, and J. R. Walters. 2004. Genetic and environmental influences on natal dispersal distance in a resident bird species. Am. Nat. 164:660-669.

Pettay, J. E., Kruuk, L. E. B., J. Jokela, and V. Lummaa. 2005. Heritability and genetic constraints of life-history evolution in preindustrial humans. Proc. Natl. Acad. Sci. 102:2838-2843.

Poissant, J., A. J. Wilson, and D. W. Coltman. 2010. Sex-specific genetic variance and the evolution of sexual dimorphism: a systematic review of cross-sex genetic correlations. Evolution 64:97–107.

Postma, E., F. Heinrich, U. Koller, R. J. Sardell, J. M. Reid, P. Arcese, and L. F. Keller. 2011. Disentangling the effect of genes, the environment and chance on sex ratio variation in a wild bird population. Proc. R. Soc. B 278:2996–3002.

Price, G. R. 1970. Selection and covariance. Nature 227:520-521.

Punzalan, D., M. Delcourt, and H. D. Rundle. 2014. Comparing the intersex genetic correlation for fitness across novel environments in the fruit fly, *Drosophila serrata*. Heredity 112:143–148.

R Core Team. 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Reid, J. M., P. Arcese, and L. F. Keller. 2006. Intrinsic parent-offspring correlation in inbreeding level in a population open to immigration. Am. Nat. 168:1-13.

Reid, J. M., P. Arcese, L. F. Keller, and S. Losdat. 2014a. Female and male genetic effects on offspring paternity: additive genetic (co)variances in female extra-pair reproduction and male paternity success in song sparrows (*Melospiza melodia*). Evolution 68:2357–2370.

Reid, J. M., P. Arcese, and S. Losdat. 2014b. Genetic covariance between components of male reproductive success: within-pair versus extra-pair paternity in song sparrows. J. Evol. Biol. 27:2046–2056.

Reid, J. M., P. Arcese, R. J. Sardell, and L. F. Keller. 2011a. Additive genetic variance, heritability and inbreeding depression in male extra-pair reproductive success. Am. Nat. 177:177-187.

Reid, J. M., P. Arcese, R. J. Sardell, and L. F. Keller. 2011b. Heritability of female extra-pair paternity rate in song sparrows (*Melospiza melodia*). Proc. R. Soc. B. 278:1114-1120.

Reid, J. M., and L. F. Keller. 2010. Correlated inbreeding among relatives: occurrence, magnitude, and implications. Evolution 64:973–985.

Reid, J. M., L. F. Keller, A. B. Marr, P. Nietlisbach, R. J. Sardell, and P. Arcese. 2014c. Pedigree error due to extra-pair reproduction substantially biases estimates of inbreeding depression. Evolution 68:802–815.

Reid, J. M., and R. J. Sardell. 2012. Indirect selection on female extra-pair reproduction? Comparing the additive genetic value of maternal half-sib extra-pair and within-pair offspring. Proc. R. Soc. B 279:1700–1708.

Reid, J. M., and M. E. Wolak. 2018. Is there indirect selection on female extra-pair reproduction through cross-sex genetic correlations with male reproductive fitness? Evol. Lett. 2:159-168.

Robertson, A. 1966. A mathematical model of the culling process in dairy cattle. Anim. Prod. 8:95–108.

Rose, M. R. 1982. Antagonistic pleiotropy, dominance, and genetic variation. Heredity 48:63–78.

Sæther, B.-E., and S. Engen. 2015. The concept of fitness in fluctuating environments. Trends Ecol. Evol. 30:273–281.

Sardell, R. J., L. F. Keller, P. Arcese, T. Bucher, and J. M. Reid. 2010. Comprehensive paternity assignment: genotype, spatial location and social status in song sparrows, *Melospiza melodia*. Mol. Ecol. 19:4352–4364.

868 Shaw, R. G. 1987. Maximum-likelihood approaches applied to quantitative genetics of natural populations. Evolution 41:812–826.

870 Shaw, R. G., and J. R. Etterson. 2012. Rapid climate change and the rate of adaptation: insight from experimental quantitative genetics. New Phytol. 195:752-765.

872 Shaw, R. G., and F. H. Shaw. 2014. Quantitative genetic study of the adaptive process. Heredity 112:13–20.

874 Smith, J. N. M., L. F. Keller, A. B. Marr, and P. Arcese (eds). 2006. Conservation and Biology of Small Populations: The Song Sparrows of Mandarte Island. Oxford University Press, Oxford.

876 Stinchcombe, J. R. 2014. Cross-pollination of plants and animals: wild quantitative genetics and plant evolutionary genetics. Pp. 128–146 *in* A. Charmantier, D. Garant, and L. E. B. Kruuk,

878 eds. Quantitative Genetics in the Wild. Oxford University Press, Oxford, UK.

Tallmon, D. A., G. Luikart, and R. S. Waples. 2004. The alluring simplicity and complex reality of

880 genetic rescue. Trends Ecol. Evol. 19:489-496.

Teplitsky, C., J. A. Mills, J. W. Yarrall, and J. Merilä. 2009. Heritability of fitness components in a

882 wild bird population. Evolution 63:716-726.

Trask, A. E., E. M. Bignal, D. I. McCracken, P. Monaghan, S. B. Piertney, and J. M. Reid. 2016.

884 Evidence of the phenotypic expression of a lethal recessive allele under inbreeding in a wild population of conservation concern. J. Anim. Ecol. 85:879–891.

886 Travisano, M., and R. G. Shaw. 2013. Lost in the map. Evolution 67:305–314.

Wagenius, S., H. H. Hangelbroek, C. E. Ridley, and R. G. Shaw. 2010. Biparental inbreeding and

888 interremnant mating in a perennial prairie plant: fitness consequences for progeny in their first eight years. Evolution 64:761-771.

890 Walling, C. A., M. B. Morrissey, K. Foerster, T. H. Clutton-Brock, J. M. Pemberton, and L. E. B. Kruuk. 2014. A multivariate analysis of genetic constraints to life history evolution in a wild

892 population of red deer. Genetics 198:1735–1749.

Walsh, B., and M. W. Blows. 2009. Abundant genetic variation + strong selection = multivariate genetic constraints: a geometric view of adaptation. Annu. Rev. Ecol. Evol. Syst. 40:41–59.

Wheelwright, N. T., L. F. Keller, and E. Postma. 2014. The effect of trait type and strength of selection on heritability and evolvability in an island bird population. Evolution 68:3325-3336.

Wilson, S., D. R. Norris, A. G. Wilson, and P. Arcese. 2007. Breeding experience and population density affect the ability of a songbird to respond to future climate variation. Proc. R. Soc. B 274:2539–2545.

Wolak, M. E. 2012. nadiv: an R package to create relatedness matrices for estimating non-additive genetic variances in animal models. Methods Ecol. Evol. 3:792–796.

Wolak, M. E., P. Arcese, L. F. Keller, P. Nietlisbach, and J. M. Reid. 2018. Data from: Sex-specific additive genetic variances and correlations for fitness in a song sparrow (*Melospiza melodia*) population subject to natural immigration and inbreeding. Dryad Digital Repository. https://doi.org/10.5061/dryad.p7p1jb3.

Wolak, M. E., and L. F. Keller. 2014. Dominance genetic variance and inbreeding in natural populations. Pp. 104–127 *in* A. Charmantier, D. Garant, and L. E. B. Kruuk, eds. Quantitative Genetics in the Wild. Oxford University Press, Oxford.

Wolak, M. E., and J. M. Reid. 2016. Is pairing with a relative heritable? Estimating female and male genetic contributions to the degree of biparental inbreeding in song sparrows (*Melospiza melodia*). Am. Nat. 187:736–752.

Wolak, M. E., and J. M. Reid. 2017. Accounting for genetic differences among unknown parents in microevolutionary studies: how to include genetic groups in quantitative genetic animal models. J. Anim. Ecol. 86:7–20.

Wolak, M. E., D. A. Roff, and D. J. Fairbairn. 2015. Are we underestimating the genetic variances of dimorphic traits? Ecol. Evol. 5:590–597.

Wolf, J. B., and M. J. Wade. 2001. On the assignment of fitness to parents and offspring: whose fitness is it and when does it matter? J. Evol. Biol. 14:347–356.

Zhang, X.-S. 2012. Fisher's geometrical model of fitness landscape and variance in fitness within a changing environment. Evolution 66:2350–2368.

Zietsch, B. P., R. Kuja-Halkola, H. Walum, and K. J. H. Verweij. 2014. Perfect genetic correlation between number of offspring and grandoffspring in an industrialized human population. Proc. Natl. Acad. Sci. 111:1032-1036.

Supporting Information

Additional supporting information may be found in the online version of this article at the publisher's website:

Appendix S1. Literature summary

Appendix S2. Overall approach and data specifications

Appendix S3. Pedigree structure and genetic groups

Appendix S4. Details of model specification and implementation

Appendix S5. Additional details of results

Appendix S6. Prior sensitivity analysis

Appendix S7. Sex-specific juvenile survival

Appendix S8. Additional analyses of adult survival

Appendix S9. Additional analysis of adult annual survival and annual reproductive success

Tables

940

Table 1. Marginal posterior means, modes (in square brackets), and 95% credible intervals (in parentheses) for latent-scale estimates from the
942 bivariate model for female and male fitness. Within the additive genetic and cohort matrices, sex-specific variances are shown along the diagonal (bold) with cross-sex covariances (COV) and correlations (r, italics) above and below the diagonal respectively. Sex-specific residual
944 variances (V_R), latent scale heritabilities (h^2_{latent}), and slopes of regressions on individual coefficient of inbreeding (β_f) and immigrant genetic group coefficient (β_{IGG}) are also shown.

	Additive genetic matrix		**Cohort matrix**		**V_R**	**h^2_{latent}**	**β_f**	**β_{IGG}**
	Female fitness	Male fitness	Female fitness	Male fitness				
Female fitness	V_A=**2.01** **[1.56]** **(0.21, 3.93)**	COV_A=0.62 [0.42] (-0.43, 1.82)	**3.12** **[2.46]** **(0.90, 5.98)**	COV=1.46 [1.16] (0.12, 2.97)	16.58 [15.29] (12.99, 21.06)	0.09 [0.08] (0.02, 0.18)	-21.41 [-19.91] (-32.53, -11.48)	-6.27 [-5.79] (-11.36, -2.00)
Male fitness	*r_A=0.38* *[0.45]* *(-0.19, 0.94)*	V_A=**1.72** **[1.70]** **(0.13, 3.39)**	*r=0.67* *[0.79]* *(0.27, 0.98)*	**1.54** **[0.97]** **(0.38, 3.17)**	15.61 [15.60] (11.86, 19.31)	0.09 [0.08] (0.01, 0.17)	-27.86 [-25.53] (-39.47, -16.50)	-5.47 [-5.19] (-9.91, -0.71)

946

Table 2. Marginal posterior means, modes (in square brackets), and 95% credible intervals (in parentheses) for latent- and observed-scale estimates from the trivariate model for juvenile survival and adult female and male lifetime reproductive success (LRS). Within the additive genetic and cohort matrices, variances are shown along the diagonal (bold) with covariances (COV) and correlations (r, italics) above and below the diagonal respectively. Residual variances (V_R), latent-scale heritabilities (h^2_{latent}), observed-scale heritabilities ($h^2_{observed}$) and evolvabilities ($I_{A\text{-}observed}$), and slopes of regressions on individual coefficient of inbreeding (β_f) and immigrant genetic group coefficient (β_{IGG}) are also shown. $I_{A\text{-}observed}$ is not applicable (NA) for juvenile survival. Posterior modes and lower 95%CI limits that converged towards zero are reported as <0.001.

	Additive genetic matrix			**Cohort matrix**			$\mathbf{V_R}$	$\mathbf{h^2_{latent}}$	$\mathbf{h^2_{observed}}$	$\mathbf{I_{A\text{-}observed}}$	$\boldsymbol{\beta_f}$	$\boldsymbol{\beta_{IGG}}$
	Juvenile survival	Female LRS	Male LRS	Juvenile survival	Female LRS	Male LRS						
Juvenile survival	**V_A=0.23 [0.002] (<0.001, 0.50)**	COV_A=0.001 [<0.001] (-0.08, 0.08)	COV_A=-0.06 [0.01] (-0.41, 0.24)	**0.73 [0.71] (0.27, 1.28)**	COV=-0.02 [-0.001] (-0.17, 0.08)	COV=-0.11 [-0.10] (-0.43, 0.21)	1 (fixed)	0.09 [0.001] (<0.001, 0.18)	0.03 [<0.001] (<0.001, 0.05)	NA	-9.31 [-9.39] (-12.85, -5.61)	-2.58 [-2.26] (-4.26, -0.85)
Female LRS	*r_A=0.03 [-0.07] (-0.73, 0.80)*	**V_A=0.05 [0.001] (<0.001, 0.18)**	COV_A=-0.04 [<0.001] (-0.26, 0.13)	*r=-0.17 [-0.48] (-0.90, 0.63)*	**0.03 [<0.001] (<0.001, 0.11)**	COV=0.01 [<0.001] (-0.07, 0.10)	0.72 [0.71] (0.54, 0.93)	0.05 [0.001] (<0.001, 0.21)	0.03 [<0.001] (<0.001, 0.13)	0.05 [0.001] (<0.001, 0.18)	-1.89 [-2.11] (-5.34, 1.43)	0.32 [0.40] (-0.85, 1.69)
Male LRS	*r_A=-0.08 [-0.19] (-0.67, 0.58)*	*r_A=-0.11 [-0.25] (-0.83, 0.61)*	**V_A=1.19 [1.00] (0.28, 2.25)**	*r=-0.23 [-0.22] (-0.72, 0.33)*	*r=0.11 [0.06] (-0.61, 0.83)*	**0.35 [0.26] (0.02, 0.79)**	1.12 [0.99] (0.49, 1.80)	0.44 [0.38] (0.15, 0.74)	0.09 [0.09] (0.03, 0.15)	1.19 [0.99] (0.24, 2.20)	-9.00 [-9.25] (-15.04, -3.15)	-0.41 [-0.72] (-3.35, 2.44)

Table 3. Marginal posterior means, modes (in square brackets), and 95% credible intervals (in parentheses) from (A) the bivariate model for adult female and male annual reproductive success (ARS) and (B) the univariate model for adult annual survival. Within the additive genetic and year matrices for ARS, variances are shown along the diagonal (bold) with covariances (COV) and correlations (r, italics) above and below the diagonal respectively. Permanent individual (V_{PI}) and residual (V_R) variances, latent-scale heritabilities (h^2_{latent}), observed-scale heritabilities ($h^2_{observed}$) and evolvabilities ($I_{A\text{-}observed}$), and slopes of regressions on individual coefficient of inbreeding (β_f) and immigrant genetic group coefficient (β_{IGG}) are also shown. I_A is not applicable (NA) for adult annual survival. Posterior modes and lower 95%CI limits that converged towards zero are reported as <0.001.

	Additive genetic matrix		**V_{PI}**	**Year matrix**		**V_R**	**h^2_{latent}**	**$h^2_{observed}$**	**$I_{A\text{-}observed}$**	***β_f***	***β_{IGG}***
	Female ARS	Male ARS		Female ARS	Male ARS						
Female ARS	**V_A=0.02** **[<0.001]** **(<0.001, 0.05)**	COV_A=0.01 [0.001] (-0.04, 0.06)	0.01 [<0.001] (<0.001, 0.04)	**0.04** **[0.03]** **(0.01, 0.08)**	COV=0.07 [0.05] (0.01, 0.15)	0.05 [0.05] (0.04, 0.07)	0.18 [0.001] (<0.001, 0.38)	0.06 [<0.001] (<0.001, 0.15)	0.02 [<0.001] (<0.001, 0.05)	-1.29 [-1.19] (-2.82, 0.23)	0.22 [0.20] (-0.38, 0.84)
Male ARS	*r_A=0.17* *[0.18]* *(-0.57, 0.88)*	**V_A=0.16** **[0.15]** **(0.02, 0.30)**	0.05 [<0.001] (<0.001, 0.15)	*r=0.67* *[0.78]* *(0.30, 0.96)*	**0.34** **[0.25]** **(0.11, 0.64)**	0.46 [0.46] (0.34, 0.58)	0.16 [0.12] (0.03, 0.31)	0.08 [0.08] (0.01, 0.15)	0.67 [0.40] (0.01, 1.64)	-4.98 [-4.98] (-7.94, -2.07)	0.15 [0.11] (-1.18, 1.61)
Adult annual survival	0.04 [<0.001] (<0.001, 0.16)		0.37 [0.01] (<0.001, 1.44)	0.69 [0.54] (0.20, 1.31)		1 (fixed)	0.02 [<0.001] (<0.001, 0.07)	0.005 [<0.001] (<0.001, 0.02)	NA	-0.49 [0.03] (-4.81, 3.83)	0.33 [0.46] (-1.48, 2.10)

970

Figure Legends

Figure 1. Phenotypic distributions of (A) fitness, (B) adult lifetime reproductive success, and (C) adult annual reproductive success measured as the number of banded chicks attributed to each focal individual. Red and blue denote females and males respectively.

Figure 2. Marginal posterior MCMC samples (bars), kernel density estimation (solid black line), posterior mean (red dotted line), 95% credible interval limits (black dashed lines), and prior (solid blue line) for the additive genetic variances (V_A) in (A) female fitness, (B) male fitness, and (C) the cross-sex additive genetic correlation (r_A) in song sparrows. In A and B, the priors are depicted over the range of each posterior distribution, but extend to substantial positive values.

Figure 3. Marginal posterior MCMC samples (bars), kernel density estimation (solid black line), posterior mean (red dotted line), 95% credible interval limits (black dashed lines), and prior (solid blue line) for the additive genetic variances (V_A) in (A) juvenile survival, (B) adult female lifetime reproductive success (LRS), and (C) adult male LRS, and the additive genetic correlations (r_A) between (D) juvenile survival and adult female LRS, (E) juvenile survival and adult male LRS, and (F) adult female and male LRS in song sparrows. Note that axis scales vary among plots. In A-C, the priors are depicted over the range of each posterior distribution, but extend to substantial positive values.

Figure 4. Marginal posterior MCMC samples (bars), kernel density estimation (solid black line), posterior mean (red dotted line), 95% credible interval limits (black dashed lines), and prior (solid blue line) for the additive genetic variances (V_A) in (A) adult female annual reproductive success (ARS), (B) adult male ARS, and (C) the cross-sex additive genetic correlation (r_A) in song

sparrows. On A and B, x-axis scales are standardized to facilitate comparison, but the y-axis scales differ. The panel A inset shows the marginal posterior distribution for female ARS on a larger scale. In A and B, the priors are depicted over the range of each posterior distribution, but extend to substantial positive values.

Figure 5. Phenotypic distributions of age-specific survival (or mortality) for adult (A) female and (B) male song sparrows, and (C) the marginal posterior distribution for additive genetic variance (V_A) in adult annual survival. In A and B, dark and light shading indicate observations of mortality and survival respectively. In C, plot attributes are as for figures 2-4.