The quantitative genetics of fitness in a wild seabird

Maria Moiron[1,2*], Anne Charmantier[1†], Sandra Bouwhuis[2†]

[1] CEFE, Univ Montpellier, CNRS, EPHE, IRD, Univ Paul Valéry Montpellier 3, Montpellier, France

[2] Institute of Avian Research, An der Vogelwarte 21, 26386 Wilhelmshaven, Germany

*Author for correspondence: mariamoironc@gmail.com

[†] Contributed equally as last authors

Running headline: Genetics of fitness in the wild

Keywords: adaptive potential, animal model, heritability, lifetime reproductive success, latent scale

ABSTRACT

Additive genetic variance in fitness is a prerequisite for adaptive evolution, as a trait must be genetically correlated with fitness to evolve. Despite its relevance, additive genetic variance in fitness has not often been estimated in nature. Here, we investigate additive genetic variance in lifetime fitness, as well as its underlying components, in common terns (*Sterna hirundo*). Using 28 years of data comprising ca. 6000 pedigreed individuals, we find that additive genetic variances in the Zero-inflated and Poisson components of lifetime fitness were nominally zero, but estimated with high uncertainty. Similarly, additive genetic variances in adult annual reproductive success and survival did not differ from zero, but were again associated with high uncertainty. Simulations suggested that we would be able to detect additive genetic variances as low as 0.05 for the Zero-inflated component of fitness, but not for the Poisson component, although having data for more generations of birds would lead to an important increase in statistical power. As such, our study suggests heritable variance in common tern fitness to be rather low if not null, shows how studying quantitative genetics of fitness in natural populations remains challenging, and highlights the importance of maintaining long-term individual-based studies of natural populations.

INTRODUCTION

Fisher's Fundamental Theorem of Natural Selection postulates that "*the rate of increase in fitness of any organism at any time is equal to its genetic variance in fitness at that time*" (Fisher 1930). As such, additive genetic variance of fitness, being equivalent to the change in mean fitness resulting from selection, has been considered the single most useful statistic quantifying selection (Burt 1995). Genetic variation in fitness also is a prerequisite for adaptive evolution, as a trait must be genetically correlated with fitness to evolve through natural selection (Robertson 1966; Price 1970). Hence, understanding the quantitative genetics of individual variation in fitness is arguably one of the most important aims in evolutionary ecology (Burt 1995; Ellegren and Sheldon 2008; Walsh and Blows 2009; Gomulkiewicz and Shaw 2013; Shaw and Shaw 2014; Hendry et al. 2018).

While additive genetic variance is of indisputable relevance in predicting evolutionary dynamics of natural populations, other sources of genetic, yet non-additive, variance are also important. Non-additive genetic variance components such as dominance and epistasis can strongly influence the mean and variance in fitness (Roff and Emerson 2006; Carroll 2007). Indeed, epistasis is known to be a key reservoir for additive genetic variance, as it can be readily converted to additive genetic variance as a result of inbreeding or environmental change (Cheverud and Routman 1995; Whitlock et al. 1995; Buskirk and Willi 2006), ultimately contributing to the evolutionary dynamics of fitness. Under scenarios of strong local adaptation and stable conditions, genetic variance is predicted to be reduced, because stabilizing selection acts to maintain locally adapted phenotypes and alleles with larger average effect would be driven close to fixation and thus contribute less to the total genetic variance (Johnson and Barton 2005). However, populations under directional selection and in variable environments will readily recruit new additive genetic variance due to recombination (Otto and Lenormand 2002). Altogether, the existence and

magnitude of genetic variance within populations will greatly differ depending on the existence and magnitude of local adaptation and/or stabilizing selection, because differences in selection regime will lead to different likelihoods of evolutionary change.

Considerable debate has surrounded the question of whether additive genetic variation in fitness should be low or not (e.g., Jones 1987; Burt 1995; Houle et al. (1996), Merilä and Sheldon 1999; Shaw and Shaw 2014), and particularly, under which conditions (e.g., Cheverud and Routman 1995; Whitlock et al. 1995), however, empirical estimates of additive genetic variance in fitness from wild populations have so far not shed the much-needed light on this debate. A recent review of 30 studies on humans, other animals and plants found that there were very few estimates of additive genetic variance (V_A) in fitness (or fitness components) in the wild, and that those that were available varied substantially, with many estimates close to zero, and few large estimates (Hendry et al. 2018). To provide some examples: Kruuk et al. (2000) found zero V_A in lifetime fitness in female Scottish red deer (*Cervus elaphus*), but some evidence for V_A in males ($V_{A\ males}$ = 0.434, SE = 0.681) (but see Foerster et al. 2007). Additive genetic variance was also estimated to be zero in both sexes of bighorn sheep (*Ovis Canadensis)* in Canada (Coltman et al. 2005) and very small in North American red squirrels (*Tamiasciurus hudsonicus*) ($V_A \sim 0$, 95% = 5.2E-07 - 1.1, McFarlane et al. 2014, see also McFarlane et al. 2015). In birds, Gustafsson (1986) estimated a zero heritability of lifetime reproductive success in male and female collared flycatchers (*Ficedula albicollis*) in Sweden. In a later study from the same population, however, Merilä and Sheldon (2000) found additive genetic variance in lifetime reproductive success for females and males to be non-zero (h^2= 0.21 ± 0.06 for females and 0.07± 0.06 for males, see also Brommer et al. (2007)). In female and male British great tits (*Parus major*), McCleery et al. (2004) found close-to-zero heritability of lifetime reproductive success (h^2= 0.00 ± 0.04 for females and

0.02± 0.04 for males). Along the same lines, Wheelwright et al. (2014) found a zero heritability for lifetime reproductive success in female savannah sparrows (*Passerculus sandwichensis*) in Canada, while Teplitsky et al. (2009) found non-zero genetic variance in lifetime reproductive success for females (h^2= 0.36 ± 0.29) and zero variance for males in a natural population of red-billed gulls (*Laurus novaehollandiae*) in New Zealand. Finally, de Villemereuil et al. (2019) showed that hihis (*Notiomystis cincta*) in New Zealand had negligible additive genetic variance in lifetime fitness, while Wolak et al. (2018) found that the song sparrows (*Melospiza melodia*) of Mandarte island in Canada harbored substantial additive genetic variance in female and male fitness ($V_{A\ female}$=2.01, 95% CI =0.21, 3.93; $V_{A\ male}$=1.72, 95% CI =0.27, 3.39).

Data constraints might partially explain the paucity of studies testing for the heritability of fitness in the wild and the heterogeneity among estimates of additive genetic variance, although steadily increasing datasets collected from long-term study populations gradually alleviate the problem (Clutton-Brock and Sheldon 2010). This increased data availability was recently accompanied by the development of (i) statistical tools designed to deal with the non-Gaussian distributions that often characterize fitness data (de Villemereuil et al. 2016; de Villemereuil 2018), as well as (ii) theoretical frameworks that facilitate the evolutionary inference of quantitative genetic parameters based on these data distributions (Morrissey and Bonnet 2019). Fitness components often follow a non-Gaussian error distribution as a result from the temporal sequence of survival (Binary data) and reproductive events (Poisson data). Modelling them accordingly, by applying generalized linear models (de Villemereuil et al. 2016; de Villemereuil 2018; Bonnet et al. 2019), offers an added benefit. Parameter estimates from a model with a Poisson error distribution fitted to absolute fitness data readily inform about the increase in fitness within a generation, while back-transformed estimates on the observed data scale inform about the increase

in fitness between generations (Morrissey and Bonnet 2019). As such, estimates of the additive genetic variance for absolute fitness on the latent scale data are equivalent to evolvability estimates (i.e., the additive genetic variance in relative fitness) on the data scale for relative fitness, and, therefore, provide evidence for Fisher's rate of evolution (Hansen et al. 2011; de Villemereuil et al. 2016). To date, only four studies have modelled the quantitative genetics of fitness in wild populations assuming a Poisson (McFarlane et al. 2014, 2015; Wolak et al. 2018) or a Zero-Inflated Poisson distribution (de Villemereuil et al. 2019).

Here, we present phenotypic and pedigree data obtained from a 28-year individual-based study on common terns (*Sterna hirundo*). The common tern is a Nearctic and Palearctic colonially breeding, serially monogamous and migratory seabird. The study colony is located in the north of Germany; common terns from this colony spend their winters in western Africa and return to the breeding colony in early spring to breed or prospect potential breeding locations (Becker and Ludwigs 2004). Common terns breed annually, both parents incubate and feed the chicks, and extra-pair paternity is rare (González-Solís et al. 2001; Becker and Ludwigs 2004). Applying a series of "animal models" to data from almost 6000 pedigreed individuals across five generations, we investigate additive genetic variance for lifetime fitness (assessed as the total number of fledglings produced by a locally-born fledgling), and two of its underlying components: annual reproductive success and annual adult survival.

METHODS

Study System

Fitness and pedigree data were collected between 1992 and 2019 as part of a long-term study of a common tern population located at the Banter See on the German North Sea coast (53°36´N,

08°06´E). The Banter See colony consists of six concrete islands, each of which is surrounded by a 60-cm wall. Walls are equipped with 44 elevated platforms, each containing an antenna which reads transponder codes. The individual-based study at the Banter See was initiated in 1992, when 101 adult birds were caught and marked with individually numbered subcutaneously injected transponders. Since 1992, all locally hatched birds are similarly marked with a transponder shortly before fledging and the presence and reproductive performance of marked individuals is monitored following a standard protocol (Becker and Wendeln 1997). As part of this protocol, the colony is checked for new clutches every 2–3 days throughout the breeding season (Zhang et al. 2015). Parents are identified using portable antennae placed around each nest for 1–2 days during incubation, which is shared by both partners. Pairs can rear up to three chicks per brood (mean successful brood size 0.41 ± 0.65 SD chicks), and can produce replacement clutches after loss of eggs or chicks. Second clutches are extremely rare (Becker and Zhang 2011).

Fitness Data

Fitness data have been collected since 1992, with data up to 2019 being available for the analyses reported here. Our initial data selection included individuals that fledged between 1992 and 2016, because previous work showed that 97% of fledglings, if they returned, did so within the first 3 years (Vedder and Bouwhuis 2018). Although we cannot directly quantify juvenile dispersal, our data suggest that that it is relatively infrequent. This is because of (i) a relatively high local return rate (28% of chicks fledged between 1992 and 2016 recruited at the colony), and (ii) rare reporting of external recruits (between 1992 and 2016, 32 fledglings from the Banter See were observed a total of 105 times in other European breeding colonies), although this reporting may not be extensive. In addition, although we cannot directly observe an individual's death, we can reliably

assume it, because adult breeders at the Banter See are highly site-faithful, evidenced by the resighting probability of individuals that bred at least once being close to one (Szostek and Becker 2012), and 96% of breeders not skipping recording by the antenna system for two or more consecutive years after first reproduction (Bouwhuis et al. 2015; Zhang et al. 2015). Based on this knowledge, we removed all birds that were observed in 2018 and/or 2019 *and* were younger than 11 years old, because (i) they are known to not be, or cannot yet be assumed to be, dead, and (ii) lifetime fitness of individuals older than 10 years and those dead showed a high correlation ($r > 0.8$) in our dataset. Hence, we included birds that have completed their life histories ($n = 5836$), as well as birds that were still alive but older than 10 years ($n = 163$) to avoid introducing a cohort truncation bias by non-randomly removing longer-lived birds (Hadfield 2008; Morrissey et al. 2012). To control for any potential confounding effect, we modelled whether an individual was considered dead or alive as a fixed effect (see below).

We quantified lifetime fitness as the total number of local fledglings that a locally-hatched fledgling produced during its lifetime. In total, our data comprise the fitness of 5999 locally-hatched fledglings (Fig. 1A). Lifetime fitness can be decomposed in two major components: juvenile survival and adult lifetime reproductive success. Juvenile survival captures survival from fledgling to adulthood (first year of breeding, once birds are adults), whereas adult lifetime reproductive success captures adult survival and reproductive success across life. These two fitness components, juvenile survival and adult lifetime reproductive success, correspond to the two mechanisms captured by the Zero-inflated Poisson distribution of lifetime fitness, and hence, we did not model them in separate analyses. However, we further decomposed adult lifetime reproductive success (LRS) into its two components: annual reproductive success (ARS) and annual adult survival (AAS). ARS was measured as the number of fledglings that an individual

produced each year between its first year of life and last registration, assigning zeroes for years of skipped reproduction or registration, and for years prior to recruitment (Fig. 1B); AAS was adult survival (1/0) to the following breeding season, measured every year from an individual's first year of life to last registration. In total, our data comprised 836 individuals with 6873 observations for ARS and AAS.

Pedigree

The pedigree was constructed by assigning all fledged offspring to their social parents, then pruned to remove individuals who were either not phenotyped or not ancestors to phenotyped individuals. For the purpose of this study, the pedigree comprised 6273 records. The maximum depth was five generations, the number of paternities and maternities 2509 and 2414, respectively. The numbers of full, maternal and paternal siblingships were 2566, 10180 and 9718, respectively. This social pedigree is a good approximation of the genetic pedigree, because common terns exhibit very low levels of extra-pair paternity (González-Solís et al. 2001).

Quantitative Genetic Models

We applied an animal model approach that combines the phenotypic information on individual fitness with information from the social pedigree (Kruuk 2004). As such, we fitted a series of univariate animal models where fitness, or one of its components, was the response variable.

To model lifetime fitness, we fitted a univariate animal model with a Zero-Inflated Poisson error distribution. We fitted a Zero-Inflated Poisson distribution to better capture the nature of our metric of lifetime fitness. Zero-inflation is often the result of two different processes involved in producing the data, i.e., the process that determines whether an event occurs or not differs from

187 the process that determines how many times an event occurs, if it does occur. In this case, a Zero-
188 Inflated Poisson model can explicitly model the two different processes, as opposed to a Poisson
189 model that assumes only a single process to be generating the data (Korner-Nievergelt et al. 2015).
190 We fitted random intercepts for individual identity linked to the pairwise relatedness matrix and
191 for hatch-year (to account for cohort effects; e.g., Vedder and Bouwhuis 2018). Because we
192 modeled lifetime fitness with a Zero-Inflated over-dispersed Poisson distribution, this model has
193 a Zero-Inflated and a Poisson component, thereby allowing the estimation of the covariance
194 between the two components for each random effect. However, a model including additive genetic
195 and hatch-year correlations between the Zero-Inflated and Poisson components of the trait did not
196 provide a better fit to the data, hence we do not report such correlations. The main models also did
197 not control for shared environmental effects between siblings (maternal, paternal, or brood effects)
198 because we did not have information on parental identity for all individuals (maternal identities =
199 2382 and paternal identities = 2481; 1271 individuals have both maternal and paternal identities
200 known), and because most fledglings came from broods where only a single individual had
201 successfully fledged (3027 broods fledged one chick, 1145 broods two, 226 broods 3, while 4
202 individuals could not be assigned to a brood). However, we did explore the potential effects of a
203 shared environment (due to maternal, paternal effects, or brood effects) by running two additional
204 animal models in which included one or two shared environmental effects as a random effect. We
205 found that there was no major influence on our estimate of additive genetic variance in fitness,
206 neither when modelling paternal and maternal effects nor when modelling brood effects, as
207 expected given that the model presented in the main text showed very low (close to or zero)
208 estimate of additive genetic variance (see Suppl. Material, Tables S1 and S2).

As fixed effects, we modelled the trait intercept and whether the individual was alive or dead (categorical variable with two levels). Additionally, we performed data simulations to investigate (i) whether we can effectively detect *small, but substantial* heritabilities and evolvabilities (*sensu* de Villemereuil et al. 2019) given our data and pedigree structure, and (ii) the improvement of our statistical power to detect small additive genetic variances in both components of lifetime fitness when the dataset and pedigree would increase in size and depth (Supplementary Material, Figs. S1-S5).

To model ARS, we assumed a Poisson error distribution with a log link function and checked whether the trait was underdispersed, which was not the case. We fitted random intercepts for individual identity linked to the pairwise relatedness matrix, individual identity not linked to the pedigree (to account for permanent environmental effects) and year of observation (to account for temporal variation across years). As fixed effects, we modelled the trait intercept and age (continuous trait ranging from 1 to 23 years), as fledgling production is known to linearly increase with age (Zhang et al. 2015) (but see Supplementary Materials, Table S3, for results of the same animal model run without age effects).

To model AAS, we assumed a binary error distribution with a logit link function and fixed the residual variance to one. We fitted random intercepts for individual identity linked to the pairwise relatedness matrix, individual identity not linked to the pedigree (to account for permanent environmental effects) and year of observation (to account for temporal variation across years). As fixed effects, we modelled the trait intercept and age (continuous trait ranging from 1 to 23 years), as AAS is known to linearly decrease with age (Zhang et al. 2015; Vedder et al. 2021) (but see Supplementary Materials, Table S3, for results of the same model ran without age effects). All quantitative genetic models were fitted using a Bayesian framework implemented in the

statistical software R (v. 3.6.1, R Core Team 2019) using the R-packages *MCMCglmm* (Hadfield 2010) and *QGglmm* (de Villemereuil et al. 2016). Heritabilities (h^2) were conditional on the variance explained by fixed effects and estimated as the proportion of the total phenotypic variance explained by the additive genetic variance. Evolvabilities (I_A) were estimated by dividing the additive genetic variance by the squared population mean (Houle 1992; Hansen et al. 2011).

For all models we used parameter-expanded priors (Hadfield 2010). We fitted different priors for each fitness component (see Supplementary Material). The number of iterations and thinning intervals were chosen for each model so as to ensure that the minimum MCMC effective sample size for all parameters was 1000. Burn-in was set to a minimum of 5000 iterations. The retained effective sample sizes yielded absolute autocorrelation values <0.1 and satisfied convergence criteria based on the Heidelberger and Welch convergence diagnostic (Heidelberger and Welch 1981). We drew inferences from the marginal posterior mode and 95% credible intervals (95% CI). Variance parameters were estimated on latent scales. To facilitate evolutionary inference (Bonnet et al. 2019; Morrissey and Bonnet 2019), we back-transformed the latent-scale posterior distributions of the quantitative genetic parameters to the observed data-scale (de Villemereuil et al. 2016).

RESULTS

Quantitative Genetics of Lifetime Fitness

Among the 5999 common tern chicks that fledged between 1992 and 2016, lifetime fitness ranged between 0 and 29 fledglings (Fig. 1A). 5231 (87.19%) fledglings obtained zero fitness, such that the distribution of fitness was strongly Zero-Inflated (Fig. 1A).

Raw mean fitness was 0.72 ± 2.52 SD fledglings. Although this would indicate the population to be in overall decline (a mean lifetime breeding success of two fledglings would be

required for the population to be stable), population size actually varied dramatically across years but was not in overall decline (Fig. S6), partially because there was a substantial influx of non-locally hatched breeders that immigrated into the population (ca. 74% ± 1 of breeders was estimated to be immigrant in any given year between 1992 and 2020). Since we do not capture or mark immigrants, we can quantify the proportion of immigrants present in our colony in a given year but we cannot include them in the pedigree or our individual-based models.

Simulations showed that, given our data structure and pedigree, we would not be able to detect what might be considered a *small, but substantial* signal for the Zero-inflated component of lifetime fitness: we generated a Zero-inflated component of fitness with an additive genetic variance of 0.01, and found that the average posterior mode was similar to the simulated value of V_A (average = 0.012 across the 100 simulations, Fig. S1), but that the lower 95% CI limit was nominally zero in most of the simulations (95% CI = 0 – 0.023, Fig. S1). When we simulated larger values of additive genetic variance (i.e., V_A = 0.05 or 0.1), our simulations showed that we would be able to detect those variance values (average = 0.053 and 95% CI = 0.028 – 0.083 across the 100 simulations for a simulated value of 0.05, and average = 0.102 and 95% CI = 0.064 – 0.145 for a simulated value of 0.1, Figs. S3 and S4).

Our quantitative genetic analysis of empirical data suggested that the additive genetic variance in the Zero-Inflated component of fitness was not different from zero, as the posterior mode of the additive genetic variance was very close to, and the lower 95% CI limit leaning towards, zero (Table 1, Fig. 2A-C). Taken together, our combination of analyses of empirical and simulated data therefore suggested there to be low (lower than 0.05) to null additive genetic variance in the Zero-inflated component of lifetime fitness, but that we lack power to determine with higher precision whether such variance is nominally zero or non-zero but very small.

279 The results for the Poisson component of lifetime fitness are less straightforward.
280 Simulations showed that, given our data structure and pedigree, we would not be able to detect
281 either *small, but substantial* or larger signals for the Poisson component of fitness: we generated a
282 Poisson-component of fitness with a series of evolvability values (I_A = 0.00, 0.01, 0.05 and 0.1),
283 and found that the lower 95% CI limit was leaning towards zero in all cases (Fig. S1-4). Our
284 analysis of the empirical data suggested the additive genetic variance of the Poisson component to
285 not differ from zero, given that the associated lower 95% CI limits of V_A, h^2 and I_A converged
286 towards zero (Table 1, Fig. 2D-F). Altogether, the combination of empirical analyses and data
287 simulations showed that we lack power to determine where the additive genetic variance in the
288 Poisson component of lifetime fitness falls within a rather large range of values (between "larger
289 than 0.1" and zero).
290 Finally, data simulations of a larger dataset with a deeper pedigree structure indicated that
291 increasing our study to include four more generations of pedigreed individuals would lead to an
292 important increase in statistical power, so that we would be able to detect additive genetic variances
293 of at least 0.05 in both components of lifetime fitness. Estimated values of additive genetic
294 variance were of similar magnitude to that of the simulated value (average posterior mode of 0.05
295 across the 100 simulations for both components of lifetime fitness), with associated 95% CI not
296 leaning towards zero in any of the cases (95% CI = 0.031- 0.064 for Zero-Inflated component, and
297 95% CI = 0.009 -0.197 for Poisson component, Fig. S5).
298
299 **Quantitative Genetics of Annual Fitness Components**
300 We investigated the Annual Reproductive Success and Annual Adult Survival of 793 fledglings
301 that survived to adulthood and bred in our population (Table 2). Raw mean annual reproductive

success was 0.70 ± 0.81 SD with a maximum of three fledglings (Fig. 1B). The posterior distribution of V_A for ARS converged toward zero (Table 2, Fig. 4A-C), suggesting that V_A is not different from zero. Raw mean annual adult survival probability was 0.85 ± 0.36 SD. The posterior modes of all quantitative genetic parameters for AAS were very close to zero (Table 2, Fig. 3A-C), with the lower 95% CI limit of all parameter estimates converging towards zero, again suggesting that V_A in AAS is not different from zero.

DISCUSSION

The most direct measure of the adaptive potential of a population is its standing additive genetic variance in fitness (Fisher 1930). Here, we estimated additive genetic variances in lifetime fitness and two of its key components in a wild colony of common terns. On the one hand, our empirical findings indicated no evidence for substantial (or different than zero) additive genetic variance in lifetime fitness, annual adult survival or annual reproductive success in this population. On the other hand, data simulations demonstrated an overall lack of statistical power to detect *small, but substantial* signals (i.e., V_A = 0.01), although statistical power differed between the two components of lifetime fitness: we would have power to detect slightly larger signals (additive genetic variances of, at least, 0.05) for the Zero-inflated, but not Poisson, component of fitness. As such, our work demonstrated that estimating additive genetic variance in fitness still is very difficult in wild populations, partly due to the expected low values of genetic variation in fitness in populations locally adapted or under stabilizing selection, and partly due to the challenges associated with collecting sufficient phenotypic and pedigreed data.

Quantitative Genetics of Lifetime Fitness

In this study, we aimed at providing much-needed empirical estimates for key quantitative genetic parameters that have rarely been estimated in the wild, and did so by applying non-Gaussian models to estimate variation in fitness (Bonnet et al. 2019; Morrissey and Bonnet 2019). Quantitative genetic parameters drawn from Poisson models can be readily interpreted in terms of evolutionary significance without back-transformation. When traits follow a log-link function, estimates of additive genetic variance for absolute fitness on the latent scale are equivalent to evolvability estimates directly on the data scale for relative fitness, and therefore, they provide evidence for Fisher's rate of evolution (Hansen et al. 2011; de Villemereuil et al. 2016; Morrissey and Bonnet 2019).

There have been around 30 studies testing for additive genetic variance in fitness in the wild (see Introduction), with, to our knowledge, only four using non-Gaussian animal models (McFarlane et al. 2014, 2015; Wolak et al. 2018; de Villemereuil et al. 2019), and only one testing for variance components of fitness using a Zero-Inflated Poisson distribution (de Villemereuil et al. 2019). Our estimate of the additive genetic variance for the Zero-inflated component of common tern lifetime fitness on the observed data-scale was nominally zero, with the lower 95% CI limit leaning towards zero (posterior mode $V_{A\ data\text{-}scale} = 0.004$, 95% CI = 0 - 0.008, Table 1), similarly to results for the hihi (de Villemereuil et al. 2019) (posterior mode $V_{A\ data\text{-}scale} \sim 0$, 95% CI = 1.4×10^{-11} - 0.0038). For the Poisson component, de Villemereuil et al. (2019) found a posterior mode of 0.0078 (95% CI = 2.3×10^{-10} - 5.7). Our posterior mode estimate was overall larger (posterior mode $V_{A\ data\text{-}scale} = 2.29$, Table 1) but associated with high uncertainty (95% CI = 0.002 - 12.3), such that the estimates from both studies remain qualitatively similar.

Given that our estimates of additive genetic variance in fitness showed very low or nominally zero values, our study implies that the adaptive potential of this natural population of

common terns will be extremely limited, although the actual potential remains partially unknown as our estimates were associated with high uncertainty. Moreover, it is important to note that we could only investigate the evolutionary potential of local recruits, as we did not have phenotypic and pedigree data to investigate the evolutionary potential of the total colony. This was because the studied population had a substantial influx of immigrants that recruited at the colony but those immigrants remained unknown across the years.

Quantitative Genetics of Fitness Components

Additive genetic variance in lifetime fitness can theoretically be decomposed into the additive genetic variances in its underlying components. The two primary components of our measure of lifetime fitness are juvenile survival and adult lifetime reproductive success. Our zero-inflation in lifetime fitness is mainly due to low juvenile survival (i.e., 74% of fledglings did not locally recruit), while the Poisson process generating the observed fitness distribution is mostly capturing adult lifetime reproductive success. If we compare our nominally zero additive genetic variance in the Zero-inflated component of lifetime fitness (Table 1) with estimates from other studies that tested for additive genetic variance in juvenile survival, we observe some differences. For instance, the study of Wolak et al. (2018) on the song sparrow population of Mandarte Island reported evidence for non-zero V_A for juvenile survival. The natural history of common terns and song sparrows differs in many ways, yet one reason for this disparity could be a difference in emigration rates, since the Mandarte population is isolated with very little juvenile emigration (Reid et al. 2021).

Adult lifetime reproductive success is the sum of annual reproductive events across the life of an individual, and hence, can be decomposed into annual reproductive success and annual adult

survival. Given the lack of substantial additive genetic variance for adult annual survival or annual reproductive success (Table 2), the decomposition of adult lifetime reproductive success into its components was not very insightful in identifying what was the most likely mechanism underlying genetic variation in adult lifetime reproductive success. This finding again contrasts with one from Mandarte's song sparrows, where quantitative genetic analyses demonstrated moderate levels of V_A in ARS (especially for males) and close to zero V_A in AAS, indicating that heritable ARS was the primary component of heritable adult LRS in that population (Wolak et al. 2018).

Limitations of studying quantitative genetics of fitness in the wild

Despite the fundamental relevance of additive genetic variance in fitness in the context of understanding adaptation and evolutionary potential, Hendry et al. (2018) found that there were very few estimates of additive genetic variance for fitness in the wild, and that those available estimates were heterogeneous, with many estimates close to zero, and very few large estimates (e.g., Gustafsson 1986; Kruuk et al. 2000; Merilä and Sheldon 2000; Coltman et al. 2005; McFarlane et al. 2014).

Data constraints might partially explain the paucity of studies testing for the heritability of fitness in the wild. Animal models are data-hungry and rely on high quality pedigree information. Researchers therefore are faced with the challenge of collecting hard-to-quantify lifetime fitness data from an unbiased sample of the population (i.e., avoiding a "missing fraction" bias) that comprises a sufficiently large number of individuals of known relatedness (Burt 1995; Merilä and Sheldon 1999; Hendry et al. 2018). In addition, even when a large dataset and pedigree are available, additive genetic variance in fitness is often expected to be low, for instance, when populations are locally adapted or under stabilizing selection, such that the power to detect small,

close to zero, additive genetic variation in fitness may be low as well. Non-zero but non-significant estimates and zero estimates might simply represent bounded estimates (i.e., when the estimated parameters are very close to zero, i.e., the lower limit of the distribution, models often fail to estimate very low values with higher precision). As pointed out by Burt (1995): "it is very difficult to get an estimate that is statistically distinguishable from zero, and the sample sizes required to do so might easily lead to despair". In light of the multiple constraints posed by data requirements and expected low values, negative results with respect to additive genetic variation in fitness should be taken and discussed with care. Simulations aimed at determining the statistical power of a given dataset and pedigree structure will help to distinguish a true negative result from a zero parameter estimated with high uncertainty (e.g., de Villemereuil et al. 2019). Overall, the field of quantitative genetics in the wild needs more and better estimates stemming from a broad taxonomic range, and to systematically associate those empirical estimates with simulations to assess the power of the dataset.

Additionally, our knowledge of the genetic architecture of fitness and fitness components is currently still limited. Extending our genomic understanding of variation in fitness in wild populations will bring important insights into how natural selection maintains genetic variation underpinning fitness, and overall will help to better predict the evolutionary dynamics of natural populations (Merilä and Sheldon 1999; Mackay 2001; Huang and Mackay 2016). Despite the clear benefits, genomic research based on quantitative trait loci (QTL) approaches or genome-wide association studies on quantitative fitness-related traits applied to natural populations has been a challenge (Slate 2004; Slate et al. 2010; Jensen et al. 2014). This initial paucity in genomic studies on quantitative traits in pedigreed natural populations was partially due to the low power to detect QTL, for instance because studies suffered from low-density linkage maps and/or relatively few

417 genotyped individuals. Nowadays, the use of powerful next-generation genomic techniques,
418 however, allows to increase the power in such studies.

419 A better understanding of the genetic architecture of fitness will also provide added
420 benefits, as, for instance, it would allow a deeper understanding of the genetic underpinnings of
421 complex traits such as fitness which might be subjected to different pleiotropic effects (Mackay
422 2001). Indeed, antagonistic pleiotropy is often assumed to underlie the commonly-observed
423 negative phenotypic correlation between the two main fitness components of lifetime fitness:
424 survival and reproductive success (also observed in the terns: Vedder et al., 2021). The
425 implications of antagonistic pleiotropy for the study of genetic variance in fitness are manifold,
426 varying according to the direction, magnitude and symmetry of the different allele effects. For
427 instance, in a simple scenario where (i) the alleles with positive effects on reproductive output
428 would have negative effects on survival, (ii) the effects on both fitness components are reciprocal
429 and (iii) symmetrical, the genetic variance in survival and reproduction would be approximately
430 the same, but the genetic variance in lifetime fitness would be very low.

431

432 **Conclusion**

433 Our quantitative genetic study of fitness in a wild population of common terns reported low to zero
434 estimates of additive genetic variance in lifetime fitness and two underlying components, which
435 were at the same time associated with high uncertainty. Those analyses, however, were
436 overshadowed by a lack of statistical power to detect additive genetic variation in fitness more
437 accurately and precisely. The continuation of long-term individual-based studies should be
438 safeguarded (also see Clutton-Brock and Sheldon 2010), such that the maturation of long-term
439 studies will offer improved opportunities for testing genetic variation in natural populations,

which, thanks to the recent development of appropriate statistical and theoretical frameworks (de Villemereuil et al. 2016; Bonnet et al. 2019; Morrissey and Bonnet 2019), will help to improve our understanding of the genetics of fitness in the wild. Ultimately, a robust quantification of the standing additive genetic variation for fitness will inform us about the rate of adaptation of populations between and within generations, and allow a better understanding of their viability in the face of the deleterious environmental effects that current climate and global changes pose.

AUTHOR CONTRIBUTIONS

M.M. conceived the study with input from S.B. and A.C. M.M. designed and conducted the analyses, and wrote the manuscript. S.B. manages the tern data and collated the dataset. All authors contributed to editing the final paper.

ACKNOWLEDGMENTS

We are indebted to Peter H. Becker for setting up the long-term common tern study and thank the numerous field workers that have contributed to compiling the data. We also thank Tim Janicke for providing enlightening discussions during manuscript conception and preparation, Timothée Bonnet for helping with the interpretation of model parameters and Andrew McAdam, Gabriel Marroig, Michael Morrissey, Jarrod Hadfield and two anonymous reviewers for their constructive feedback on our manuscript. M.M. was funded by a Marie Curie Individual Fellowship (PLASTIC TERN, Grant Agreement Number 793550) and an Alexander von Humboldt postdoctoral fellowship. The authors declare no conflict of interest.

DATA ARCHIVING

Data will be archived in the Dryad Digital Repository upon acceptance.

REFERENCES

Becker, P. H., and J.-D. L. Ludwigs. 2004. Sterna hirundo Common Tern. Pp. 93–139 *in* Parkin D, ed. The Birds of the Western Palearctic Update. Oxford University Press, Oxford, UK.

Becker, P. H., and H. Wendeln. 1997. A New Application for Transponders in Population Ecology of the Common Tern. Condor 99:534–538.

Becker, P. H., and H. Zhang. 2011. Renesting of Common Terns Sterna hirundo in the life history perspective. J. Ornithol. 152:213–225.

Bonnet, T., M. B. Morrissey, and L. E. B. Kruuk. 2019. Estimation of Genetic Variance in Fitness, and Inference of Adaptation, When Fitness Follows a Log-Normal Distribution. J. Hered. 110:393–395.

Bouwhuis, S., O. Vedder, and P. H. Becker. 2015. Sex-specific pathways of parental age effects on offspring lifetime reproductive success in a long-lived seabird. Evolution (N. Y). 69:1760–1771.

Brommer, J. E., M. Kirkpatrick, A. Qvamström, and L. Gustafsson. 2007. The intersexual genetic correlation for lifetime fitness in the wild and its implications for sexual selection. PLoS One 2:1–6.

Burt, A. 1995. The evolution of fitness. Evolution. Evolution (N. Y). 49:1–8.

Buskirk, J. Van, and Y. Willi. 2006. The change in quantitative genetic variation with inbreeding. Evolution (N. Y). 60:2428–2434.

Carroll, S. P. 2007. Brave New World: the epistatic foundations of natives adapting to invaders. Genetica 129:193–204.

Cheverud, J. M., and E. J. Routman. 1995. Epistasis and its contribution to genetic variance components. Genetics 139:1455–1461.

Clutton-Brock, T., and B. C. Sheldon. 2010. Individuals and populations: the role of long-term, individual-based studies of animals in ecology and evolutionary biology. Trends Ecol. Evol. 25:562–573.

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

de Villemereuil, P. 2018. Quantitative genetic methods depending on the nature of the phenotypic trait. Ann. N. Y. Acad. Sci. 1422:29–47.

de Villemereuil, P., A. Rutschmann, K. D. Lee, J. G. Ewen, P. Brekke, and A. W. Santure. 2019. Little Adaptive Potential in a Threatened Passerine Bird. Curr. Biol. 29:889-894.e3.

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.

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

Fisher, R. 1930. The Genetical Theory of natural Selection. Oxford.

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 447:1107–1110.

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

González-Solís, J., E. Sokolov, and P. H. Becker. 2001. Courtship feedings, copulations and paternity in common terns, Sterna hirundo. Anim. Behav. 61:1125–1132.

Gustafsson, L. 1986. Lifetime Reproductive Success and Heritability: Empirical Support for

Fisher's Fundamental Theorem. Am. Nat. 128:761–764.

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

Hadfield, J. D. 2010. MCMC Methods for Multi-response Generalized Linear Mixed Models : The MCMCglmm R Package. J. Stat. Softw. 33:1–22.

Hansen, T. F., C. Pélabon, and D. Houle. 2011. Heritability is not Evolvability. Evol. Biol. 38:258–277.

Heidelberger, P., and P. D. Welch. 1981. A spectral method for confidence interval generation and run length control in simulations. Commun. ACM 24:233–245.

Hendry, A. P., D. J. Schoen, M. E. Wolak, and J. M. Reid. 2018. The Contemporary Evolution of Fitness. Annu. Rev. Ecol. Evol. Syst. 49:457–476.

Houle, D. 1992. Comparing evolvability and variability of quantitative traits. Genetics 130:195–204.

Houle, D., B. Morikawa, and M. Lynch. 1996. Comparing mutational variabilities. Genetics 143.

Huang, W., and T. F. C. Mackay. 2016. The Genetic Architecture of Quantitative Traits Cannot Be Inferred from Variance Component Analysis. PLOS Genet. 12:e1006421.

Jensen, H., M. Szulkin, and J. Slate. 2014. Molecular quantitative genetics. Pp. 209–227 *in* Quantitative Genetics in the Wild. Oxford University Press.

Johnson, T., and N. Barton. 2005. Theoretical models of selection and mutation on quantitative traits. Philos. Trans. R. Soc. B Biol. Sci. 360:1411–1425.

Jones, J. S. 1987. The heritability of fitness: Bad news for 'good genes'? Trends Ecol. Evol. 2:35–38.

Korner-Nievergelt, F., T. Roth, S. von Felten, J. Guélat, B. Almasi, and P. Korner-Nievergelt.

2015. Bayesian data analysis in ecology using linear models with R, BUGS, and Stan.

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

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. 97:698–703.

Mackay, T. F. C. 2001. The Genetic Architecture of Quantitative Traits. Annu. Rev. Genet. 35:303–339.

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.

McFarlane, S. E., J. C. Gorrell, D. W. Coltman, M. M. Humphries, S. Boutin, and A. G. McAdam. 2015. The nature of nurture in a wild mammal’s fitness. Proc. R. Soc. B Biol. Sci. 282:20142422.

Merilä, J., and B. C. Sheldon. 1999. Genetic architecture of fitness and nonfitness traits: Empirical patterns and development of ideas.

Merilä, J., and B. C. Sheldon. 2000. Lifetime reproductive success and heritability in nature. Am. Nat. 155:307–310.

Morrissey, M. B., and T. Bonnet. 2019. Analogues of the fundamental and secondary theorems of selection, assuming a log-normal distribution of expected fitness. J. Hered. 110:396–402.

556 Morrissey, M. B., D. J. Parker, P. Korsten, J. M. Pemberton, L. E. B. Kruuk, and A. J. Wilson.
557 2012. The prediction of adaptive evolution: Empirical application of the secondary theorem
558 of selection and comparison to the breeder's equation. Evolution (N. Y). 66.
559 Otto, S. P., and T. Lenormand. 2002. Resolving the paradox of sex and recombination. Nat. Rev.
560 Genet. 3:252–261.
561 Price, G. R. 1970. Selection and Covariance. Nature 227:520–521.
562 R Core Team. 2019. A language and environment for statistical computing. R Found. Stat.
563 Comput. Vienna:Austria.
564 Reid, J. M., P. Arcese, P. Nietlisbach, M. E. Wolak, S. Muff, L. Dickel, and L. F. Keller. 2021.
565 Immigration counter-acts local micro-evolution of a major fitness component: Migration-
566 selection balance in free-living song sparrows. Evol. Lett. 5:48–60.
567 Robertson, A. 1966. A mathematical model of the culling process in dairy cattle. Anim. Sci.
568 8:95–108.
569 Roff, D. A., and K. Emerson. 2006. Epistasis and dominance: evidence for differential effects in
570 life-history versus morphological traits. Evolution (N. Y). 60:1981–1990.
571 Shaw, R. G., and F. H. Shaw. 2014. Quantitative genetic study of the adaptive process. Heredity
572 (Edinb). 112:13–20.
573 Slate, J. 2004. Quantitative trait locus mapping in natural populations: progress, caveats and
574 future directions. Mol. Ecol. 14:363–379.
575 Slate, J., A. W. Santure, P. G. D. Feulner, E. A. Brown, A. D. Ball, S. E. Johnston, and J.
576 Gratten. 2010. Genome mapping in intensively studied wild vertebrate populations. Trends
577 Genet. 26:275–284.
578 Szostek, K. L., and P. H. Becker. 2012. Terns in trouble: demographic consequences of low

breeding success and recruitment on a common tern population in the German Wadden Sea. J. Ornithol. 153:313–326.

Teplitsky, C., J. A. Mills, J. W. Yarrall, and J. Merilä. 2009. Heritability of fitness components in a wild bird population. Evolution (N. Y). 63:716–726.

Vedder, O., and S. Bouwhuis. 2018. Heterogeneity in individual quality in birds: overall patterns and insights from a study on common terns. Oikos 127:719–727.

Vedder, O., I. Pen, and S. Bouwhuis. 2021. How fitness consequences of early-life conditions vary with age in a long-lived seabird: A Bayesian multivariate analysis of age-specific reproductive values. J. Anim. Ecol. 1365-2656.13471.

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 (N. Y). 68:3325–3336.

Whitlock, M. C., P. C. Phillips, F. B. G. Moore, and S. J. Tonsor. 1995. Multiple fitness peaks and epistasis. Annu. Rev. Ecol. Syst. 26:601–629.

Wolak, M. E., P. Arcese, L. F. Keller, P. Nietlisbach, and J. M. Reid. 2018. Sex-specific additive genetic variances and correlations for fitness in a song sparrow (Melospiza melodia) population subject to natural immigration and inbreeding. Evolution (N. Y). 72:2057–2075.

Zhang, H., O. Vedder, P. H. Becker, and S. Bouwhuis. 2015. Age-dependent trait variation: the relative contribution of within-individual change, selective appearance and disappearance in a long-lived seabird. J. Anim. Ecol. 84:797–807.

FIGURES

Figure 1. Phenotypic distributions of A) lifetime fitness measured as the total number of fledglings a locally-hatched fledgling produced in its lifetime (with the inset showing the distribution for non-zero fitness in more detail), and B) annual reproductive success, measured as the number of fledglings an adult breeder produced in a year.

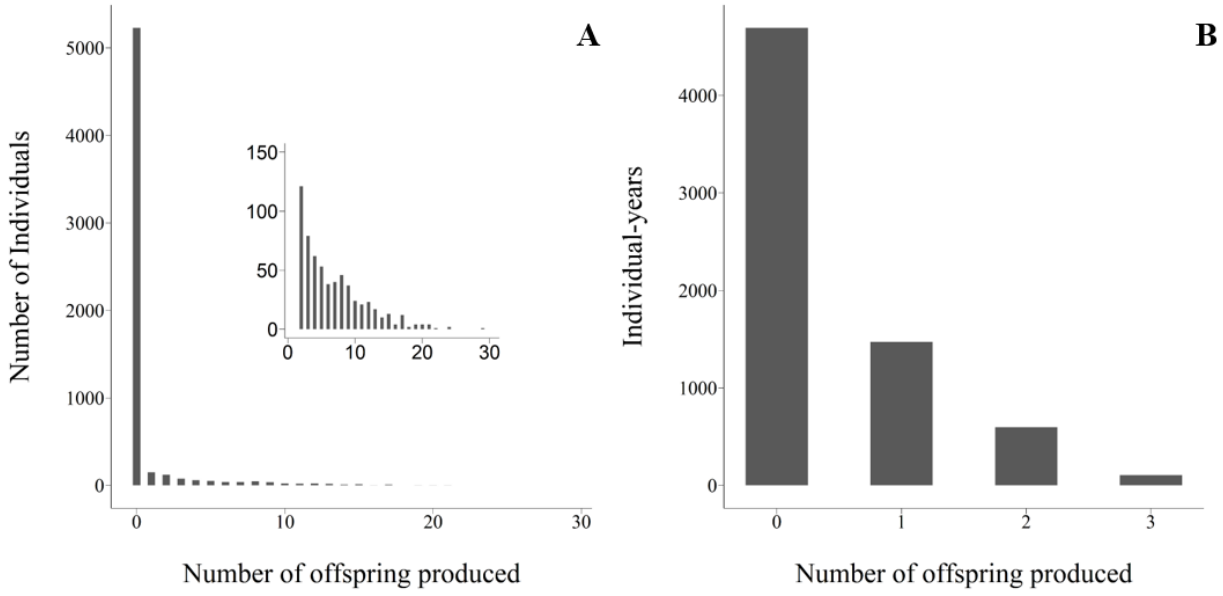


Figure 2. Posterior MCMC samples (bars), kernel density estimation (solid black line), posterior mean (red dotted line), 95% Credible Intervals (black dashed lines) and prior (solid blue line) for the A) additive genetic variance (V_A), B) heritability (h^2) and C) evolvability (I_A) of the Zero–Inflated component of lifetime fitness, and the D) additive genetic variance (V_A), E) heritability (h^2) and F) evolvability (I_A) of the Poisson component of lifetime fitness. Distributions are reported on the observed data scale.

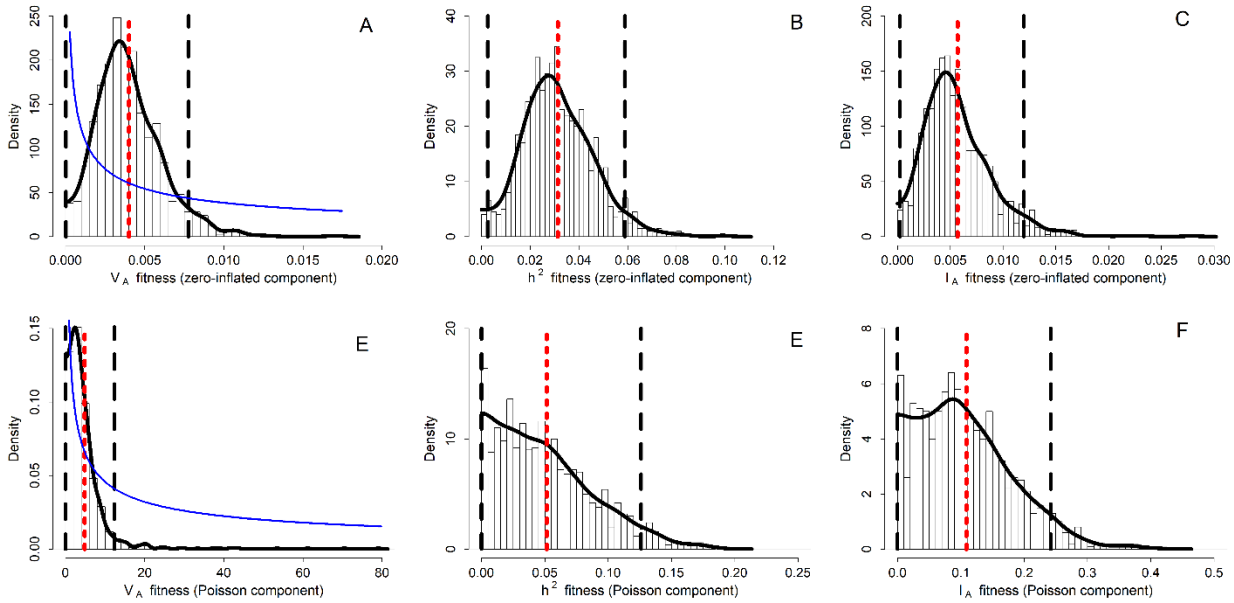


Figure 3. Posterior MCMC samples (bars), kernel density estimation (solid black line), posterior mean (red dotted line), 95% Credible Intervals (black dashed lines) and prior (solid blue line) for the A) additive genetic variance (V_A), B) heritability (h^2) and C) evolvability (I_A) of annual adult survival (AAS). Distributions are reported on the observed data scale.

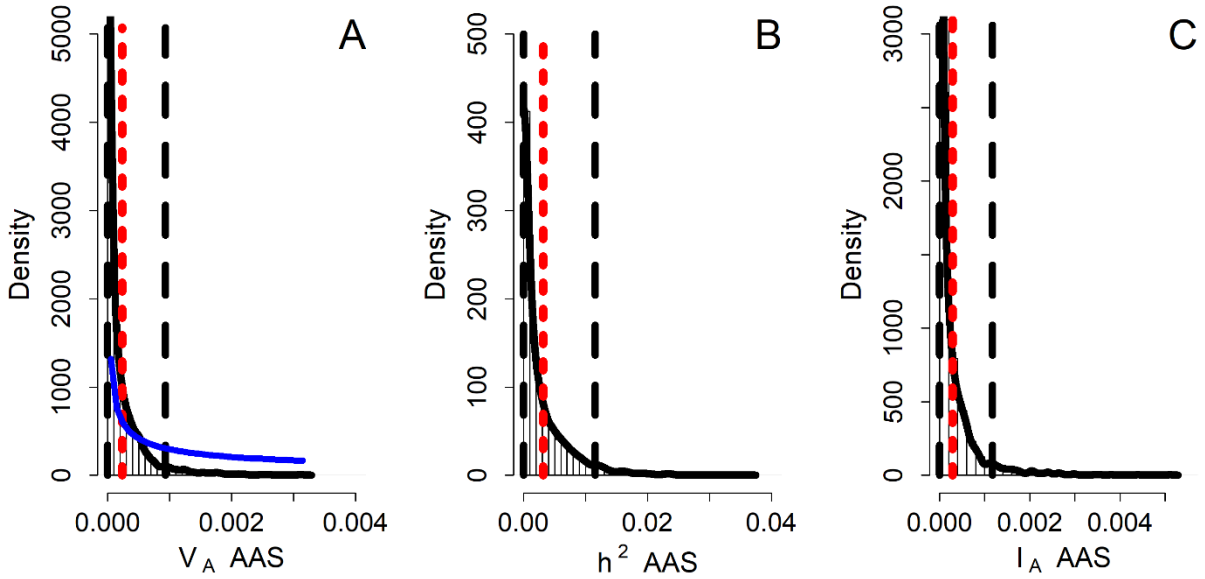
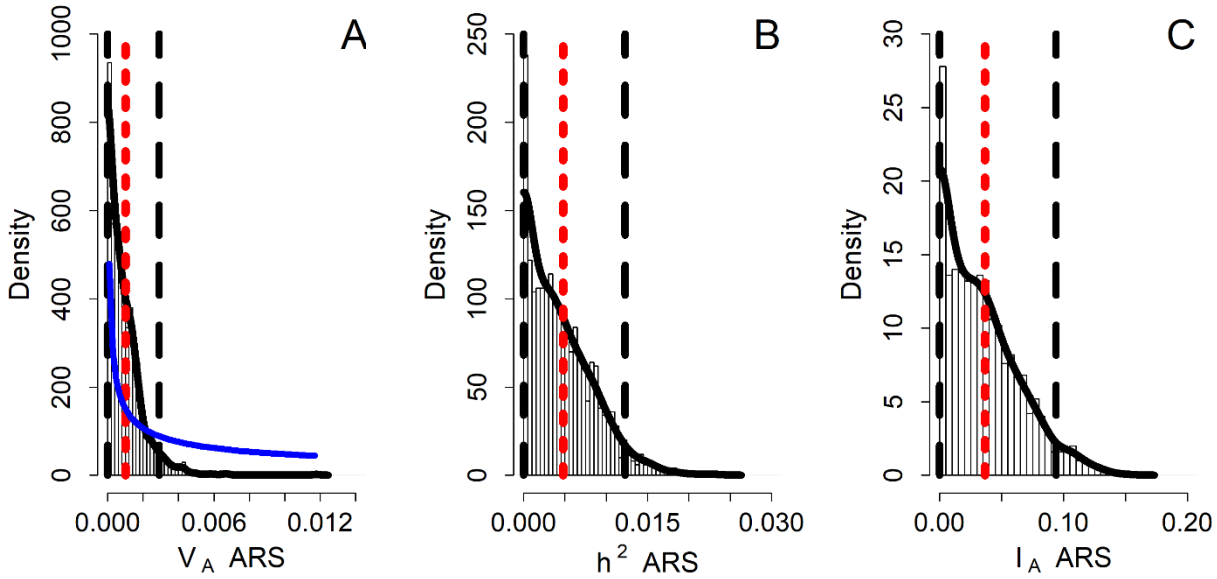


Figure 4. Posterior MCMC samples (bars), kernel density estimation (solid black line), posterior mean (red dotted line), 95% Credible Intervals (black dashed lines) and prior (solid blue line) for the A) additive genetic variance (V_A), B) heritability (h^2) and C) evolvability (I_A) of annual reproductive success (ARS). Distributions are reported on the observed data scale.



TABLES

Table 1. Posterior modes and 95% Credible Intervals (in brackets) for observed data-scale variance estimates from quantitative genetic analyses of lifetime fitness.

Model component	Number of individuals	Pop. Mean	V_P	V_A	h^2	I_A
Zero-inflated	5999	0.854 (0.777,0.908)	0.119 (0.083,0.173)	0.00401 (0,0.008)	0.031 (0.003,0.059)	0.006 (0,0.012)
Poisson		5.71 (3.86,10.2)	17.2 (20.4,549)	2.29 (0.002,12.3)	0.0232 (0,0.126)	0.0878 (0,0.242)

The results are shown for the Zero-inflated and Poisson components of the model. All statistics (Pop. Mean, population mean; V_P, phenotypic variance; V_A, additive genetic variance; h^2, heritability; I_A, evolvability) presented in the table are reported on the observed data-scale.

642 **Table 2.** Posterior modes and 95% Credible Intervals (in brackets) for observed data-scale variance estimates from quantitative genetic
643 analyses of annual reproductive success (ARS) and annual adult survival (AAS).

Fitness component	Sample size	Number of individuals	Pop. Mean	V_P	V_A	h^2	I_A
ASS	6873	836	0.940003 (0.855951,0.97266)	0.056445 (0.028845,0.125507)	0.000006 (0,0.00093)	0.000095 (0,0.011505)	0.000006 (0,0.001173)
ARS			0.142 (0.108,0.236)	0.157 (0.115,0.365)	0.000 (0,0.003)	0.000 (0,0.012)	0.000 (0,0.094)

644 All statistics (Pop. Mean, population mean; V_P, phenotypic variance; V_A, additive genetic variance; h^2, heritability; I_A, evolvability)
645 presented in the table are reported on the observed data scale.