The quantitative genetics of fitness in a wild seabird

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

[1] Centre d'Ecologie Fonctionnelle et Evolutive, Univ Montpellier, CNRS, EPHE, IRD, 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

This preprint was updated after three rounds of peer-review and published in Evolution.

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 and annual fitness 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 effectively 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, for which adequate statistical power would require c. two more decades (four tern generations) of data collection. As such, our study suggests heritable variance in common tern fitness to be rather low if not zero, shows how studying the quantitative genetics of fitness in natural populations remains challenging, and highlights the importance of maintaining long-term individual-based studies of natural populations.

Keywords: adaptive potential, additive genetic variance, heritability, lifetime reproductive success, log-normal fitness

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 in 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).

Considerable debate has surrounded the question of whether or not additive genetic variation in fitness is expected be low (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). Empirical estimates of additive genetic variance in fitness from wild populations are relatively scarce (e.g., Gustafsson 1986; Kruuk et al. 2000; Merilä and Sheldon 2000; McCleery et al. 2004; Coltman et al. 2005; Brommer et al. 2007; Foerster et al. 2007; Teplitsky et al. 2009; Wheelwright et al. 2014; McFarlane et al. 2014, 2015; Wolak et al. 2018; de Villemereuil et al. 2019), and have so far not shed much light on this debate, since estimates vary substantially, with many estimates close to zero, and few large estimates (review by Hendry et al. 2018). Overall, Hendry and colleagues (2018) tentatively concluded that the evolvability of fitness (measured as the square of the coefficient of additive genetic variance in fitness) is usually less than 0.2.

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 growing 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). To date, only four studies have modelled the quantitative genetics of fitness in wild populations assuming a non-Gaussian distribution (McFarlane et al. 2014, 2015; Wolak et al. 2018; de Villemereuil et al. 2019). Additive genetic variance in fitness was estimated to be very small in North American red squirrels (*Tamiasciurus hudsonicus*) ($V_A \sim 0$, 95% = 5.2 x 10^{-07} - 1.1, McFarlane et al. 2014, see also McFarlane et al. 2015). In birds, de Villemereuil et al. (2019) showed that hihis (*Notiomystis cincta*) in New Zealand had negligible additive genetic variance in lifetime fitness ($V_{A\ \text{Zero-Inflated component}} \sim 0$, 95% CI = 1.4 x 10^{-11} - 0.0038 and $V_{A\ \text{Poisson component}}$ =0.0078, 95% CI = 2.3 x 10^{-10} - 5.7), 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 ($V_{A\ \text{female}}$ = 2.01, 95% CI = 0.21 - 3.93; $V_{A\ \text{male}}$ = 1.72, 95% CI =0.27 - 3.39).

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 annual components: annual reproductive success and adult annual 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

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). Since there is little standardized monitoring in areas around the focal colony, we cannot directly quantify juvenile dispersal. However, we do know that there is (i) a relatively high local return rate (26% of chicks fledged between 1992 and 2016 returned to the colony, of which 14% recruited), and (ii) only 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). 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 a resighting probability of breeding individuals 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 number of local fledglings that a locally-hatched fledgling produced during its lifetime, for a total of 5999 locally-hatched fledglings (Fig. 1A) and decomposed it into two major components: juvenile survival and adult lifetime reproductive success. Juvenile survival captures survival from fledgling to age 1, inferred from whether a fledgling became a local recruit in later years, whereas adult lifetime reproductive success captures adult survival and reproductive success from age 1 onwards. These two fitness components correspond to the two mechanisms captured by the Zero-inflated Poisson distribution of lifetime fitness. We further decomposed adult lifetime reproductive success into its two components: annual reproductive success (ARS) and adult annual survival (AAS). ARS was measured as the number of fledglings that an individual produced each year between age 1 and last registration, assigning zeroes for years of skipped reproduction or registration, and for years prior to recruitment (Fig. 1B). Similarly, AAS was adult survival (1/0) to the following breeding season, measured every year from age 1 to last registration (inferring missing direct observations prior to recruitment from later observations). 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 pruned pedigree comprised 6290 records. The maximum depth was five generations, the number of paternities and maternities 2417 and 2520, respectively. The numbers of full, paternal and maternal siblingships were 2594, 10229 and 9807, respectively (see Supplementary Material for further information on the population relatedness structure). 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 components 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 a process that determines whether an event occurs or not, which differs from the Poisson process that determines how many times an event occurs. In this case, a Zero-Inflated Poisson model can explicitly model the two different processes, as opposed to a Poisson model that assumes only a single process to be generating the data (Korner-Nievergelt et al. 2015). We fitted random intercepts for individual identity linked to the pairwise relatedness matrix and for hatch-year (to account for cohort effects; e.g., Vedder and Bouwhuis 2018). Because we modeled lifetime fitness with a Zero-Inflated over-dispersed Poisson distribution, we could estimate the covariance between the Zero-inflated and Poisson components for each variance component. However, a model including additive genetic and hatch-year covariances between the Zero-Inflated and Poisson components of the trait did not provide a better fit to the data, hence we do not model such covariances. The main models presented also did not control for shared environmental effects between siblings (maternal, paternal, or brood effects), because we did not have information on parental identity for all individuals (maternal identities =

171 2382 and paternal identities = 2481; 1271 individuals have both maternal and paternal identities
172 known, see Supplementary Material for detailed information on the population relatedness
173 structure), and because most fledglings came from broods where only a single individual had
174 successfully fledged (3027 broods fledged one chick, 1145 broods two, 226 broods 3, while 4
175 individuals could not be assigned to a brood). However, we did explore the potential effects of a
176 shared environment (due to maternal, paternal effects, or brood effects) by running two additional
177 animal models which included one or two shared environmental effects as random effect(s). We
178 found that there was no major influence on our estimate of additive genetic variance in lifetime
179 fitness components, as expected given that the model presented in the main text returned a very
180 low (close to or zero) estimate of additive genetic variance (see Suppl. Material, Tables S1 and
181 S2).

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

189 To model ARS, we assumed a Poisson error distribution with a log link function and
190 checked whether the trait was underdispersed, which was not the case. We fitted random intercepts
191 for individual identity linked to the pairwise relatedness matrix, individual identity not linked to
192 the pedigree (to account for permanent environmental effects) and year of observation (to account
193 for temporal variation across years). As fixed effects, we modelled the trait intercept and age

194 (continuous trait ranging from 1 to 23 years), as fledgling production is known to linearly increase
195 with age (Zhang et al. 2015) (but see Supplementary Materials, Table S3, for results of the same
196 animal model without age effects).

197 To model AAS, we assumed a binary error distribution with a logit link function and fixed
198 the residual variance to one. We fitted random intercepts for individual identity linked to the
199 pairwise relatedness matrix, individual identity not linked to the pedigree (to account for
200 permanent environmental effects) and year of observation (to account for temporal variation across
201 years). As fixed effects, we modelled the trait intercept and age (continuous trait ranging from 1
202 to 23 years), as AAS is known to linearly decrease with age (Zhang et al. 2015; Vedder et al. 2021)
203 (but see Supplementary Materials, Table S3, for results of the same animal model without age
204 effects).

205 All quantitative genetic models were fitted using a Bayesian framework implemented in
206 the statistical software R (v. 3.6.1, R Core Team 2019) using the R-packages *MCMCglmm*
207 (Hadfield 2010) and *QGglmm* (de Villemereuil et al. 2016). Posterior distributions were plotted
208 using the R-package *wolakR (github.com/matthewwolak/wolakR).* Narrow-sense heritabilities (h^2)
209 were conditional on the variance explained by fixed effects and were estimated as the proportion
210 of the total phenotypic variance explained by the additive genetic variance. Evolvabilities (I_A) were
211 estimated by dividing the additive genetic variance by the squared population mean (Houle 1992;
212 Hansen et al. 2011).

213 For all models we used parameter-expanded priors (Hadfield 2010). We fitted different
214 priors for each fitness component (see Supplementary Material). The number of iterations and
215 thinning intervals were chosen for each model so as to ensure that the minimum MCMC effective
216 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 posterior mode and 95% credible intervals (95% CI). 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 data-scale (de Villemereuil et al. 2016).

RESULTS

Quantitative Genetics of Lifetime Fitness Components

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 and did not 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 replicates, Fig. S1), but the lower 95% CI limit was on average zero across replicates (95% CI = 0 – 0.023 and lower 95% CI exceeded a value of 0.0001 only 72 times across the 100 replicates, 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 (average = 0.053 and 95% CI = 0.028 – 0.083 across the 100 replicates 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). Lower 95% CI always exceeded a value of 0.0001 in both simulated cases (Figs. S3 and S4).

Our quantitative genetic analysis of empirical data suggested that the additive genetic variance in the Zero-Inflated component of lifetime 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 effectively zero, or non-zero but very small.

The results for the Poisson component of lifetime fitness are less straightforward. Simulations showed that, given our data structure and pedigree, we would not be able to detect either *small, but substantial* or larger signals for the Poisson component of fitness: we generated a Poisson-component of fitness with a series of evolvability values (I_A = 0.00, 0.01, 0.05 and 0.1), and found that the lower 95% CI limit was on average zero in all cases (i.e., lower 95% CI did not exceed a value of 0.0001 in the vast majority of the 100 replicates, Fig. S1-4). Our analysis of the empirical data suggested that the additive genetic variance of the Poisson component did not differ from zero, given that the associated lower 95% CI limits of V_A, h^2 and I_A converged towards zero

(Table 1, Fig. 2D-F). Altogether, the combination of empirical analyses and data simulations showed that we lacked power to determine where the additive genetic variance in the Poisson component of lifetime fitness falls within a rather large range of values (between "larger than 0.1" and zero).

Finally, simulation of a larger dataset with a deeper pedigree structure indicated that increasing our study to include four more generations of pedigreed individuals would lead to an important increase in statistical power, such that we would be able to detect additive genetic variances of at least 0.05 in both components of lifetime fitness. Estimated values of additive genetic variance were of similar magnitude to that of the simulated value (average posterior mode of 0.05 across the 100 replicates for both components of lifetime fitness), with non-zero lower 95% CI in both cases (95% CI = 0.031- 0.064 for Zero-Inflated component, and 95% CI = 0.009 - 0.197 for Poisson component, Fig. S5).

Quantitative Genetics of Annual Fitness Components

We investigated the Annual Reproductive Success and Adult Annual Survival of 836 fledglings 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 adult annual 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 and annual fitness components in a wild colony of common terns. On the one hand, our empirical findings indicated no evidence for substantial (or different from zero) additive genetic variance in lifetime fitness components, adult annual survival or annual reproductive success. 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 is very difficult in wild populations, partly due to the expected low values of genetic variation in fitness in locally adapted populations, but also to the challenges associated with collecting sufficient phenotypic and pedigreed data.

Quantitative Genetics of Lifetime and Annual Fitness Components

There have been around 30 studies testing for additive genetic variance in fitness in the wild, 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). Our estimate of the additive genetic variance for the Zero-inflated component of lifetime fitness on the data-scale was effectively zero, with a zero lower 95% CI limit (posterior mode $V_{A\ data\text{-}scale}$ = 0.004, 95% CI = 0 - 0.008, Table 1), similarly to results for another bird species, the hihi (posterior mode $V_{A\ data\text{-}scale} \sim 0$, 95% CI = 1.4×10^{-11} - 0.0038, de Villemereuil et al. 2019). For the Poisson component, de Villemereuil et al. (2019)

found a posterior mode of 0.0078 (95% CI = 2.3 x 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 null 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 (i.e., local recruits and immigrants).

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 return to the colony), 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.

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 adult annual survival. Our quantitative genetic analyses of these two annual fitness components revealed a lack of substantial additive genetic variance for both (Table 2). This finding again contrasts with one

from Mandarte's song sparrows, where there was evidence for moderate levels of additive genetic variance in ARS (especially for males) and close to zero in AAS, indicating that heritable ARS was the primary component of heritable adult lifetime reproductive success in that population (Wolak et al. 2018).

Limitations of studying quantitative genetics of fitness in the wild

Estimating quantitative genetic parameters with precision is a data-hungry endeavor. 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 the "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 pedigree is available, additive genetic variance in fitness is often expected to be low, for instance, when populations are locally adapted, such that the power to detect small, close to zero, additive genetic variation in fitness may be low as well. 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". Our data simulations reveal that we would need at least four more generations of terns to statistically differentiate between an underpowered and a true zero estimate of additive genetic variance for the Poisson component of lifetime fitness. Increasing our pedigree by four more generations would require roughly two more decades of data collection, i.e. a non-negligible amount of funding and logistic effort. This extrapolation should, however, be taken with care, as it is challenging to predict the population dynamics for the next twenty years, and/or whether the relatedness structure of the population will increase or decrease as the rates of emigration and immigration may change with population growth (e.g., Szostek and Becker 2014). 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 discussed with caution. Nevertheless, 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).

In addition to the difficulty of estimating the heritability of fitness with precision, our knowledge of the genetic architecture of fitness components is limited. Extending our genomic understanding of fitness variation in wild populations will bring important insights into how genetic variation underpinning fitness may be maintained, 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 associations in natural populations was a challenge (Slate 2004; Slate et al. 2010; Jensen et al. 2014), partially due to the low power to detect QTL, for instance because studies suffered from low-density linkage maps and/or relatively few genotyped individuals. Nowadays, the use of powerful next-generation genomic techniques, however, allows to increase the power in such studies.

A better understanding of the genetic architecture of fitness will also provide added benefits, as, for instance, it would allow a deeper understanding of the genetic underpinnings of complex traits such as fitness, which might be subjected to different pleiotropic effects (Mackay 2001). For instance, antagonistic pleiotropy is often assumed to underlie the negative phenotypic correlation between the two main components of lifetime fitness: survival and reproductive success (also observed in the terns: Vedder et al., 2021).

Conclusion

Our quantitative genetic study of fitness in a wild population of common terns reported low to zero estimates of additive genetic variance in lifetime and annual fitness components, which were at the same time associated with high uncertainty. Those analyses, however, were overshadowed by a lack of statistical power to detect additive genetic variation in fitness more accurately and precisely. The continuation of long-term individual-based studies should be safeguarded (also see Clutton-Brock and Sheldon 2010), such that the maturation of long-term 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 in fitness will inform us about the rate of adaptation of populations, and allow a better understanding of their viability in the face of the deleterious environmental effects resulting from current climate and global changes.

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 the manuscript. M.M. was funded by a Marie Skłodowska-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 have been archived in the Dryad Digital Repository: https://doi.org/10.5061/dryad.8kprr4xqj (Moiron et al. 2022).

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.

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.

450 Gomulkiewicz, R., and R. G. Shaw. 2013. Evolutionary rescue beyond the models. Philos. Trans.
451 R. Soc. B Biol. Sci. 368.
452 González-Solís, J., E. Sokolov, and P. H. Becker. 2001. Courtship feedings, copulations and
453 paternity in common terns, Sterna hirundo. Anim. Behav. 61:1125–1132.
454 Gustafsson, L. 1986. Lifetime Reproductive Success and Heritability: Empirical Support for
455 Fisher's Fundamental Theorem. Am. Nat. 128:761–764.
456 Hadfield, J. D. 2008. Estimating evolutionary parameters when viability selection is operating.
457 Proc. R. Soc. B Biol. Sci. 275:723–734.
458 Hadfield, J. D. 2010. MCMC Methods for Multi-response Generalized Linear Mixed Models :
459 The MCMCglmm R Package. J. Stat. Softw. 33:1–22.
460 Hansen, T. F., C. Pélabon, and D. Houle. 2011. Heritability is not Evolvability. Evol. Biol.
461 38:258–277.
462 Heidelberger, P., and P. D. Welch. 1981. A spectral method for confidence interval generation
463 and run length control in simulations. Commun. ACM 24:233–245.
464 Hendry, A. P., D. J. Schoen, M. E. Wolak, and J. M. Reid. 2018. The Contemporary Evolution of
465 Fitness. Annu. Rev. Ecol. Evol. Syst. 49:457–476.
466 Houle, D. 1992. Comparing evolvability and variability of quantitative traits. Genetics 130:195–
467 204.
468 Houle, D., B. Morikawa, and M. Lynch. 1996. Comparing mutational variabilities. Genetics 143.
469 Huang, W., and T. F. C. Mackay. 2016. The Genetic Architecture of Quantitative Traits Cannot
470 Be Inferred from Variance Component Analysis. PLOS Genet. 12:e1006421.
471 Jensen, H., M. Szulkin, and J. Slate. 2014. Molecular quantitative genetics. Pp. 209–227 *in*
472 Quantitative Genetics in the Wild. Oxford University Press.

473 Jones, J. S. 1987. The heritability of fitness: Bad news for ‘good genes’? Trends Ecol. Evol.
474 2:35–38.
475 Korner-Nievergelt, F., T. Roth, S. von Felten, J. Guélat, B. Almasi, and P. Korner-Nievergelt.
476 2015. Bayesian data analysis in ecology using linear models with R, BUGS, and Stan.
477 Kruuk, L. E. B. 2004. Estimating genetic parameters in natural populations using the “animal
478 model.” Philos. Trans. R. Soc. B Biol. Sci. 359:873–890.
479 Kruuk, L. E. B., T. H. Clutton-Brock, J. Slate, J. M. Pemberton, S. Brotherstone, and F. E.
480 Guinness. 2000. Heritability of fitness in a wild mammal population. Proc. Natl. Acad. Sci.
481 97:698–703.
482 Mackay, T. F. C. 2001. The Genetic Architecture of Quantitative Traits. Annu. Rev. Genet.
483 35:303–339.
484 McCleery, R. H., R. A. Pettifor, P. Armbruster, K. Meyer, B. C. Sheldon, and C. M. Perrins.
485 2004. Components of Variance Underlying Fitness in a Natural Population of the Great Tit
486 Parus major. Am. Nat. 164:E62–E72.
487 McFarlane, S. E., J. C. Gorrell, D. W. Coltman, M. M. Humphries, S. Boutin, and A. G.
488 Mcadam. 2014. Very low levels of direct additive genetic variance in fitness and fitness
489 components in a red squirrel population. Ecol. Evol. 4:1729–1738.
490 McFarlane, S. E., J. C. Gorrell, D. W. Coltman, M. M. Humphries, S. Boutin, and A. G.
491 McAdam. 2015. The nature of nurture in a wild mammal’s fitness. Proc. R. Soc. B Biol.
492 Sci. 282:20142422.
493 Merilä, J., and B. C. Sheldon. 1999. Genetic architecture of fitness and nonfitness traits:
494 Empirical patterns and development of ideas.
495 Merilä, J., and B. C. Sheldon. 2000. Lifetime reproductive success and heritability in nature. Am.

Nat. 155:307–310.

Moiron, M., A. Charmantier, and S. Bouwhuis. 2022. Data from: The quantitative genetics of fitness in a wild seabird. Dryad, doi: doi.org/10.5061/dryad.8kprr4xqj.

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.

Morrissey, M. B., D. J. Parker, P. Korsten, J. M. Pemberton, L. E. B. Kruuk, and A. J. Wilson. 2012. The prediction of adaptive evolution: Empirical application of the secondary theorem of selection and comparison to the breeder's equation. Evolution (N. Y). 66.

Price, G. R. 1970. Selection and Covariance. Nature 227:520–521.

R Core Team. 2019. A language and environment for statistical computing. R Found. Stat. Comput. Vienna:Austria.

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

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

Slate, J. 2004. Quantitative trait locus mapping in natural populations: progress, caveats and future directions. Mol. Ecol. 14:363–379.

Slate, J., A. W. Santure, P. G. D. Feulner, E. A. Brown, A. D. Ball, S. E. Johnston, and J. Gratten. 2010. Genome mapping in intensively studied wild vertebrate populations. Trends Genet. 26:275–284.

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.

519 Teplitsky, C., J. A. Mills, J. W. Yarrall, and J. Merilä. 2009. Heritability of fitness components in
520 a wild bird population. Evolution (N. Y). 63:716–726.
521 Vedder, O., and S. Bouwhuis. 2018. Heterogeneity in individual quality in birds: overall patterns
522 and insights from a study on common terns. Oikos 127:719–727.
523 Vedder, O., I. Pen, and S. Bouwhuis. 2021. How fitness consequences of early-life conditions
524 vary with age in a long-lived seabird: A Bayesian multivariate analysis of age-specific
525 reproductive values. J. Anim. Ecol. 1365-2656.13471.
526 Walsh, B., and M. W. Blows. 2009. Abundant Genetic Variation + Strong Selection =
527 Multivariate Genetic Constraints: A Geometric View of Adaptation. Annu. Rev. Ecol. Evol.
528 Syst. 40:41–59.
529 Wheelwright, N. T., L. F. Keller, and E. Postma. 2014. The effect of trait type and strength of
530 selection on heritability and evolvability in an island bird population. Evolution (N. Y).
531 68:3325–3336.
532 Whitlock, M. C., P. C. Phillips, F. B. G. Moore, and S. J. Tonsor. 1995. Multiple fitness peaks
533 and epistasis. Annu. Rev. Ecol. Syst. 26:601–629.
534 Wolak, M. E., P. Arcese, L. F. Keller, P. Nietlisbach, and J. M. Reid. 2018. Sex-specific additive
535 genetic variances and correlations for fitness in a song sparrow (Melospiza melodia)
536 population subject to natural immigration and inbreeding. Evolution (N. Y). 72:2057–2075.
537 Zhang, H., O. Vedder, P. H. Becker, and S. Bouwhuis. 2015. Age-dependent trait variation: the
538 relative contribution of within-individual change, selective appearance and disappearance in
539 a long-lived seabird. J. Anim. Ecol. 84:797–807.

FIGURE LEGENDS

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.

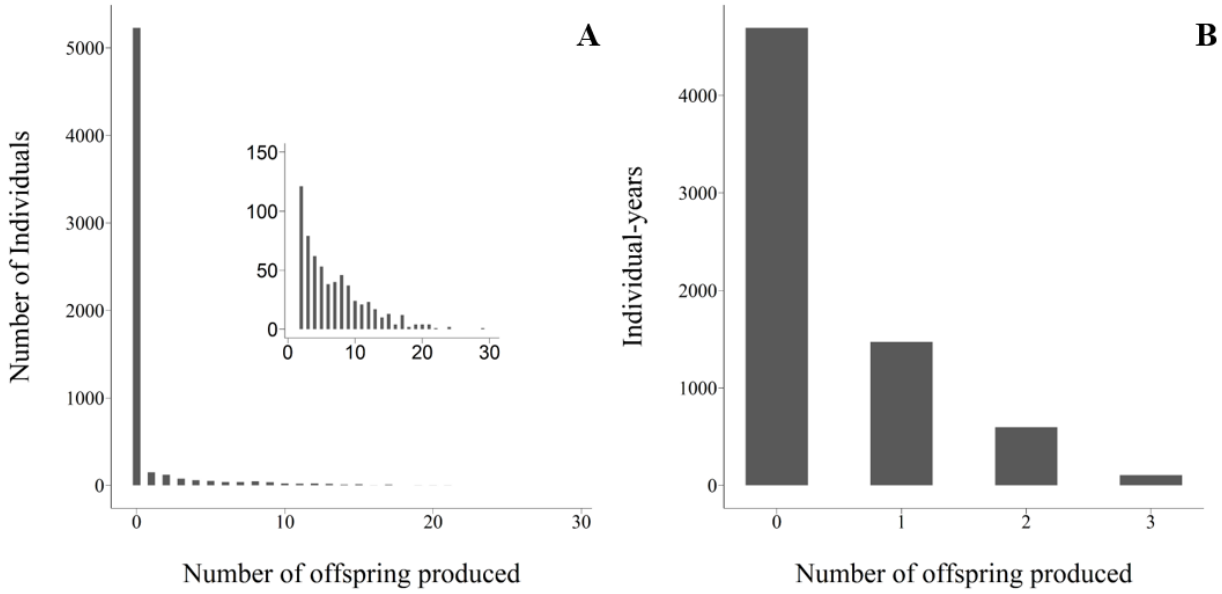


Figure 2. Posterior MCMC samples (bars), kernel density estimation (solid black line), posterior mean (red dotted line), 95% Credible Intervals (black dashed lines) and prior (solid blue line) for the A) additive genetic variance (V_A), B) heritability (h^2) and C) evolvability (I_A) of the Zero–Inflated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

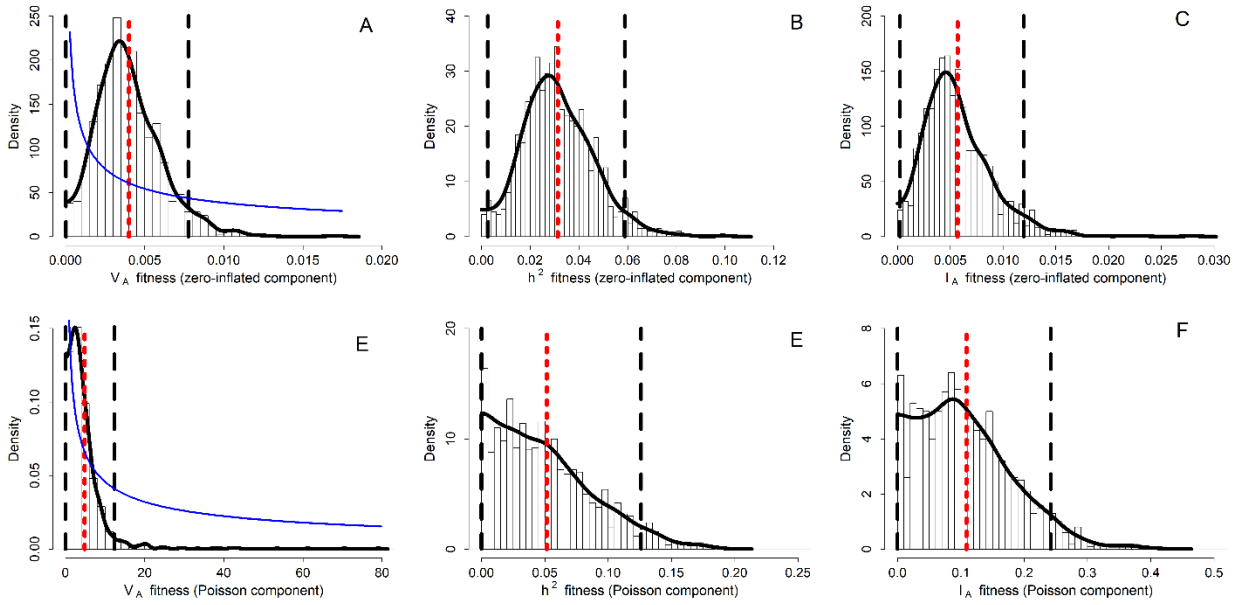
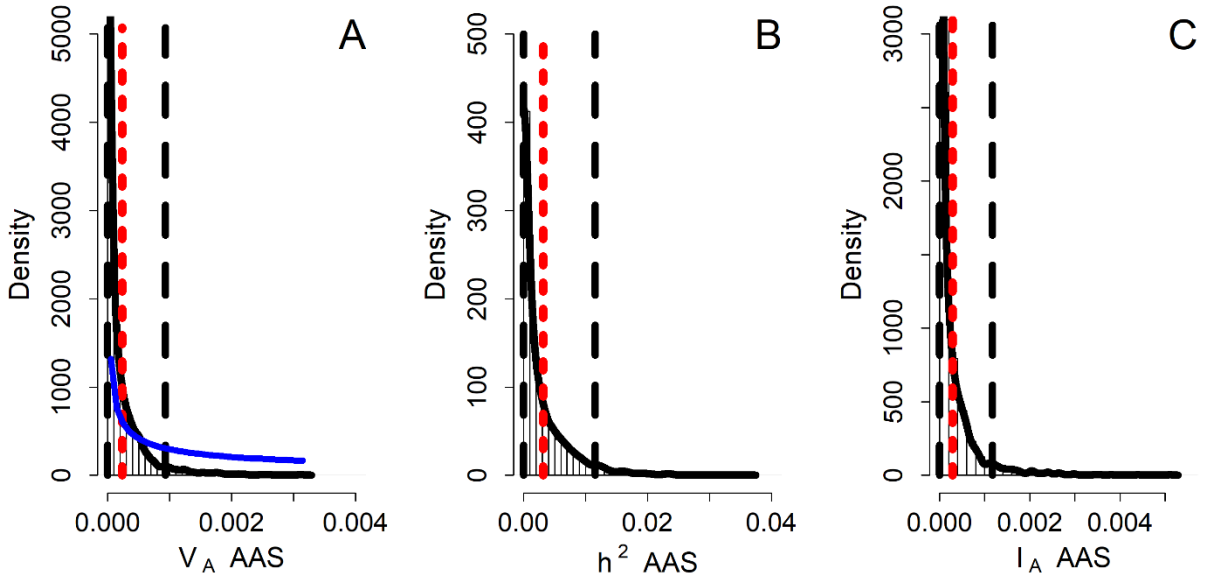
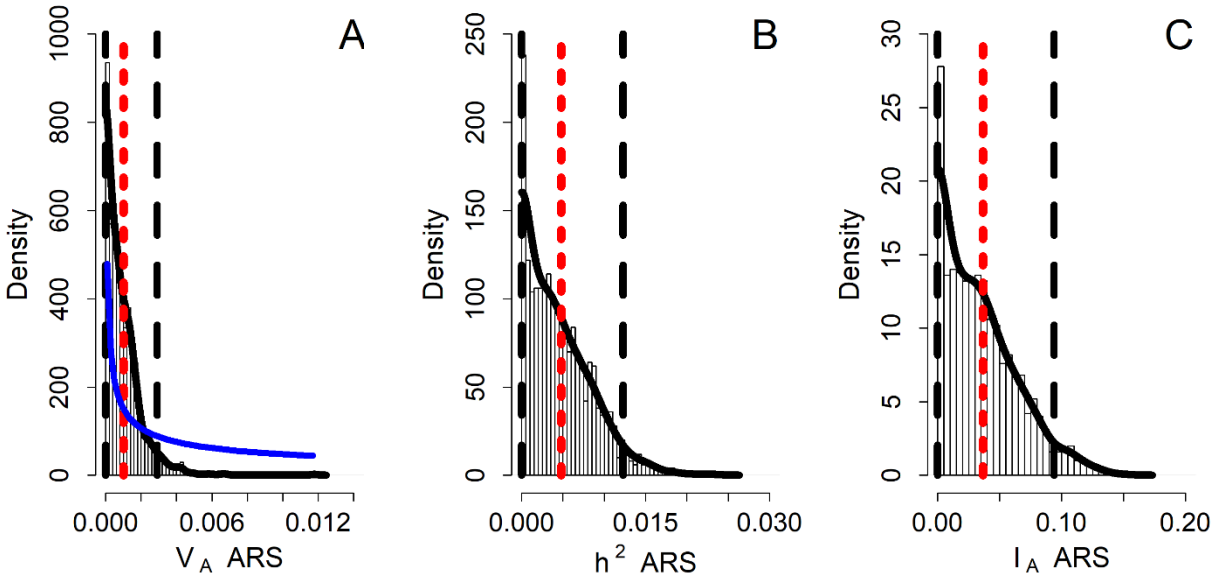


Figure 3. Posterior MCMC samples (bars), kernel density estimation (solid black line), posterior mean (red dotted line), 95% Credible Intervals (black dashed lines) and prior (solid blue line) for the A) additive genetic variance (V_A), B) heritability (h^2) and C) evolvability (I_A) of adult annual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



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

559



560

TABLES

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

Fitness component	$N_{individuals}$	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)	0.004 (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)	0.023 (0,0.126)	0.088 (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 data-scale.

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

Fitness component	$N_{observations}$	$N_{individuals}$	Pop. Mean	V_P	V_A	h^2	I_A
ASS	6873	836	0.940 (0.855,0.972)	0.056 (0.029,0.126)	0.000 (0,0.001)	0.0001 (0,0.012)	0.000 (0,0.001)
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)

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