doi:10.1111/evo.12499

The effect of trait type and strength of selection on heritability and evolvability in an island bird population

Nathaniel T. Wheelwright,[1,2] Lukas F. Keller,[3] and Erik Postma[3]

[1] *Department of Biology, Bowdoin College, Brunswick, Maine 04011*

[2] *E-mail: nwheelwr@bowdoin.edu*

[3] *Institute of Evolutionary Biology and Environmental Studies, University of Zurich, 8057 Zurich, Switzerland*

Received February 21, 2013
Accepted July 15, 2014

The heritability (h^2) of fitness traits is often low. Although this has been attributed to directional selection having eroded genetic variation in direct proportion to the strength of selection, heritability does not necessarily reflect a trait's additive genetic variance and evolutionary potential ("evolvability"). Recent studies suggest that the low h^2 of fitness traits in wild populations is caused not by a paucity of additive genetic variance (V_A) but by greater environmental or nonadditive genetic variance (V_R). We examined the relationship between h^2 and variance-standardized selection intensities (i or β_σ), and between evolvability (I_A:V_A divided by squared phenotypic trait mean) and mean-standardized selection gradients (β_μ). Using 24 years of data from an island population of Savannah sparrows, we show that, across diverse traits, h^2 declines with the strength of selection, whereas I_A and I_R (V_R divided by squared trait mean) are independent of the strength of selection. Within trait types (morphological, reproductive, life-history), h^2, I_A, and I_R are all independent of the strength of selection. This indicates that certain traits have low heritability because of increased residual variance due to the age at which they are expressed or the multiple factors influencing their expression, rather than their association with fitness.

KEY WORDS: **Animal model, fitness, Fisher's fundamental theorem, mean-standardized selection gradient, natural selection, selection intensity, Savannah sparrow.**

A major challenge in evolutionary biology is explaining variation in the evolutionary potential among traits (Houle 1992; Merilä and Sheldon 1999). Historically, narrow-sense heritability (h^2) has been used as a measure of evolutionary potential; h^2 estimates the relative importance of additive genetic variance (V_A) in shaping phenotypic variance (V_P) ($h^2 = V_A/V_P$) (Falconer and Mackay 1996). A common observation is that the phenotypic traits that have the largest influence on an individual's fitness have the lowest h^2 (Mousseau and Roff 1987; Falconer and Mackay 1996; Merilä and Sheldon 1999). The traditional explanation is that directional selection on traits important for fitness ("fitness traits") eliminates inferior alleles and fixes superior alleles, thereby limiting evolutionary potential by exhausting genetic variation in phenotypic traits in direct proportion to their effect on fitness, in accordance with Fisher's fundamental theorem (Fisher 1930; Kimura 1958; Gustafsson 1986; Falconer and Mackay 1996; Teplitsky et al. 2009). However, heritability is problematic for comparing levels of additive genetic variation of different traits. Because $h^2 = V_A/V_P$, and because V_P comprises both heritable (V_A) and nonheritable (environmental and nonadditive genetic) residual variation ($V_R = V_P - V_A$), the relatively low heritability of fitness traits could be the result of elevated V_R rather than exhausted V_A (Price and Schluter 1991; Houle 1992; Merilä and Sheldon 1999; Hansen et al. 2011).

According to the univariate breeders' equation, the evolutionary response to selection (R) equals h^2 times the selection differential (S, the covariance between a trait and relative fitness; $R = h^2S$) (Lynch and Walsh 1998; Hansen et al. 2011). Because h^2 measures the relative amount of additive genetic variation underlying a particular trait at a particular time, it only allows for

EVOLUTION
INTERNATIONAL JOURNAL OF ORGANIC EVOLUTION

a comparison of the evolutionary response of traits under equally strong selection (Postma 2014). However, by definition fitness traits are under stronger selection than nonfitness traits (Mousseau and Roff 1987; Falconer and Mackay 1996; Merilä and Sheldon 1999). Thus, h^2 by itself cannot be used to compare the evolutionary potential of different traits.

Houle (1992) introduced the concept "evolvability," which is the "expected percent change in a trait under a unit strength of selection" (Hansen et al. 2011). Evolvability is best measured as the mean-standardized additive genetic variance underlying a trait (Garcia-Gonzalez et al. 2012). The coefficient of additive genetic variation (CV_A = square root of V_A divided by the phenotypic mean of the trait, multiplied by 100) is frequently used to measure evolutionary potential (Teplitsky et al. 2009; Garcia-Gonzalez et al. 2012). However, I_A, which is V_A divided by the squared phenotypic trait mean, multiplied by 100 (Houle 1992), has a more direct evolutionary interpretation and is preferable for comparing estimates of evolutionary responses of different traits under directional selection (Hansen et al. 2011).

Fitness traits such as longevity are themselves affected by numerous physiological, morphological, and behavioral traits, each of which in turn is affected by environmental factors. As a consequence, fitness traits could be expected to have relatively high V_R because of the many possible sources of environmental variation that influence traits that are "one step further down the causal pathway from genes to phenotype" (Price and Schluter 1991). The same would be true for traits that integrate environmental influences across the entire lifespan as opposed to being expressed only at a specific age. According to this reasoning, fitness traits would be predicted to have lower h^2 than traits under weaker selection, not because of low V_A but because of high V_R (Merilä and Sheldon 1999).

On the other hand, because fitness traits are likely to have more loci affecting their expression, they present a bigger mutational target than simple (nonfitness) traits, which potentially results in faster replenishment of V_A, as originally speculated by Kimura (1958). This leads to the opposite prediction, that V_A should be *greater* in traits closely linked to fitness (Houle et al. 1996; Houle 1998; Merilä and Sheldon 1999). Despite substantial theoretical and empirical work, this issue remains unresolved (Merilä and Sheldon 1999; Teplitsky et al. 2009).

The few studies that have quantified additive genetic variance and residual variance in natural vertebrate populations have produced conflicting results. In red-billed gulls (*Larus novaehollandiae*), h^2 and CV_A declined as the trait's correlation with fitness increased (Teplisky et al. 2009). In contrast, in two populations of nest-box-breeding birds and two populations of ungulates, only h^2 was negatively correlated with fitness, whereas CV_A showed no correlation or was positively correlated with fitness (Kruuk et al. 2000; Merilä and Sheldon 2000; McCleery et al. 2004; Coltman et al. 2005).

Although groundbreaking, these field studies were limited by the small number of phenotypic traits examined, and they were often confounded by a history of human management of the study populations (e.g., provisioning of artificial nest sites for birds, culling of ungulate herds) that may have altered selection and reduced environmental and/or genetic variance (Houle 1992; Coltman et al. 2005). Moreover, studies so far have examined the relationship between a trait's correlation with fitness and h^2 or CV_A across relatively few fundamentally different types of traits (McCleery et al. 2004). However, distinct types of traits may be quite unlike in their genetic architecture (e.g., number and interactions among of underlying loci), and hence in the absolute and/or relative amount of V_A and V_R (Kimura 1958; Houle et al. 1996; Merilä and Sheldon 1999). Combining such different traits into a single broad analysis could generate an association even if at a finer scale there was no functional relationship between the strength of selection and evolvability. To properly test the hypothesis that genetic variation underlying traits reflects a balance between the elimination of alleles via directional selection and the restoration of genetic variation via mutation, an association between fitness, heritability, and variance components should be detectable *within* as well as *across* trait types. Finally, all of the field studies listed above related both h^2 and CV_A to a variance-standardized measure of the strength of selection (e.g., the unsigned correlation of a trait with fitness). However, as pointed out by Hereford et al. (2004), mean-standardized measures of genetic variation (i.e. I_A., CV_A) should be related to mean-standardized measures of selection (selection gradients = β_μ), whereas variance-standardized measures of genetic variation (i.e., h^2) should be related to variance-standardized measures of selection (selection intensity, $i = \beta_\sigma$).

Here, we explore these relationships using data gathered during a 24-year study of a wild migratory population of Savannah sparrows (*Passerculus sandwichensis*) breeding on a remote island in Canada. We measured a diversity of phenotypic traits that were categorized as related to morphology, reproductive behavior, or life history. Because this population shows extreme natal and breeding philopatry, we were able to obtain an accurate measure of lifetime production of recruits to use as comprehensive measure of fitness that incorporates both viability and fecundity (Wheelwright and Mauck 1998). Taking advantage of a pedigree covering up to 12 generations, we applied animal models to estimate quantitative genetics parameters for all traits.

Methods

STUDY SITE AND FIELD METHODS

Since 1987 NTW has studied Savannah sparrows on Kent Island, a 100-ha island in the Bay of Fundy, New Brunswick, Canada (44°35'N, 66°46'W). The study site consists of three

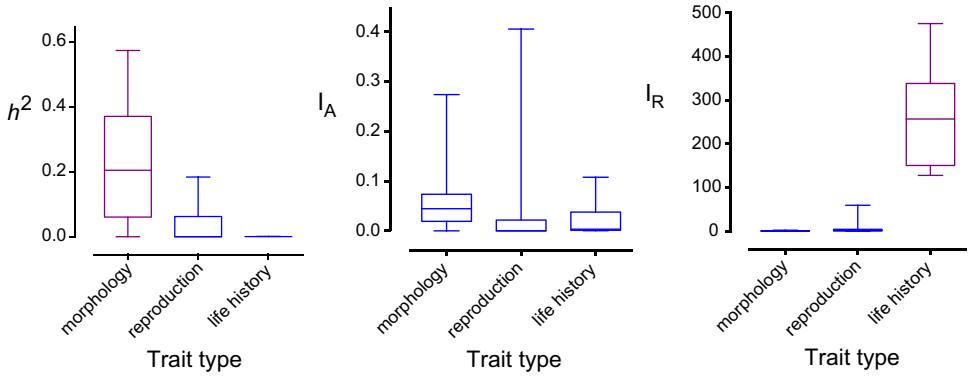

Figure 1. Box and whiskers plots showing heritability (h^2), evolvability (I_A), and mean-standardized residual variance (I_R) for three trait types ($N = 50$ morphological traits, 26 reproductive traits, 12 life history traits) in Savannah sparrows on Kent Island, New Brunswick, Canada. Horizontal line indicates median, box indicates 10th and 90th percentiles, and whiskers indicate minimum and maximum values. I_A and I_R are given as percentages (see Hansen et al. 2011).

fields totaling 10 ha in area, within which all adults are uniquely color-banded, all breeding pairs identified, all nests located, and all nestlings banded. Observations and measurements used in this study were made daily throughout the breeding season (late May–early August) from 1987 through 2005 (Wheelwright et al. 2006). The population, which has never been hunted, managed, or provisioned with food or artificial nest sites, has changed little in size, and its natural open habitat has remained relatively constant since the island became a biological field station 76 years ago. For further details of the study site and field methods, see Freeman-Gallant et al. (2003, 2006) and Wheelwright et al. (2006).

DETERMINATION OF PEDIGREES

The pedigree for this study was determined by detailed observations of uniquely banded birds. Social parents were confirmed by daily observations of mate guarding, copulations, territory defense, incubation, and nestling feeding; offspring were banded in the nest at seven days of age. Genetic paternity was confirmed in three of the years of the study using microsatellites. For further details on pedigrees, see Supplementary Information Online.

QUANTIFYING AND CLASSIFYING PHENOTYPIC TRAITS

Phenotypic traits were distinguished by sex and age class (nestling, juvenile, 1-year-old, >1-year-old), and grouped into one of three trait types (morphology, reproduction, life history) (Figs. 1 and 3; see Table 1 and Supplementary Tables S1 and S2 in Supplementary Information Online for detailed description of traits and their measurement). NTW made all morphological measurements on juveniles and adults as well as many of the other measurements and observations; measurements were highly repeatable between years (unpubl. data). For measurements made by different field assistants, we corrected for year and observer effects. Counting features measured at different ages or in different sexes as separate traits (see Hansen et al. 2011), in total we measured 88 phenotypic traits. We subsequently reanalyzed our data excluding traits highly correlated with other traits (see Methods and Supplementary Tables S1–S3). Results are given as mean and standard deviation (SD) or mean ± standard error (SE), as indicated.

To enable comparisons of absolute levels of additive genetic and residual variance between traits with different means or measured on different scales, we used I_A and I_R (see eq. 9 in Houle 1992). I_R is similar to I_A and is calculated as V_R divided by the squared trait mean and multiplied by 100. Although most previous studies have used the coefficient of residual variance, CV_R (McCleery et al. 2004; Teplitsky et al. 2009), I_R is preferable because it can be compared directly to I_A and summed with it to determine I_P (the phenotypic variance divided by the squared trait mean). Following Hansen et al. (2011, Table 1), all I_A and I_R values are expressed as percentages by multiplying them by 100. To test for biases introduced by skewed distributions and high variance relative to the mean (e.g., most life history traits), we reanalyzed a subset of our traits after applying Kleckowski's transformation (log-transformation after adding a constant), following Teplitsky et al. (2009). As found by McCleery et al. (2004), the transformation yielded qualitatively similar results to scaling to the mean (using I_A and I_R, or CV_A and CV_R) so we present only the latter.

Traits expressed at different ages by both sexes were treated as distinct traits to account for potential sex and life stage differences in genetic architecture (Jones 1987; Jensen et al. 2003; Steven et al. 2007). For example, wing length at the nestling, juvenile, yearling, and older adult stages, in males versus females, was analyzed as eight different traits. Genetic correlations likely exist

Table 1. Heritability (h^2), SE of h^2, I_A (= V_A divided by the squared phenotypic mean, multiplied by 100), SE of V_A, I_R (= V_R divided by the squared phenotypic mean, multiplied by 100), selection intensity ($i = \beta_\sigma$), and mean-standardized selection gradient (β_μ) for phenotypic traits measured on Savannah sparrows on Kent Island, New Brunswick, Canada. For comparison with earlier papers, $I_A = (CV_A)^2 \times 100$; $I_R = (CV_R)^2 \times 100$.

A. Females

Trait type	Trait	h^2	SE h^2	I_A	SE V_A	I_R	i	β_μ
(i) Nestlings								
Morphology	Condition (mass/wing)**	0.005	0.062	0.0069	0	1.2650	0.30	2.82
Morphology	Condition (mass/tarsus)**	0.056	0.112	0.0436	0	0.7297	0.17	1.96
Morphology	Mass	0.094	0.353	0.2740	0.280	1.1783	0.49	3.65
Morphology	Wing length	0	0	0.0000	1.591	2.3223	0.29	2.22
Morphology	Tarsus length	0.013	0.157	0.0060	0.290	0.4700	0.19	3.11
(ii) Juveniles								
Morphology	Condition (mass/wing)**	0.209	0.168	0.0677	0	0.2758	0.14	2.60
Morphology	Condition (mass/tarsus)**	0.137	0.275	0.0567	0	0.3555	−0.10	−1.53
Morphology	Mass	0.204	0.168	0.0633	0.150	0.2478	0.17	3.07
Morphology	Wing length	0.285	0.156	0.0161	0.384	0.0405	0.06	2.45
Morphology	Tarsus	0.473	0.233	0.0566	0.122	0.0598	0.06	1.61
Morphology	Tail length	0	0.115	0.0001	2.453	0.8697	0.62	9.88
(iii) 1-year-old								
Morphology	Condition (mass/wing)**	0.065	0.108	0.0540	0	0.7676	0.10	1.20
Morphology	Condition (mass/tarsus)**	0.087	0.118	0.0723	0	0.7697	0.13	1.48
Morphology	Mass	0.133	0.105	0.1171	0.344	0.7590	0.09	1.06
Morphology	Wing length	0.265	0.099	0.0163	0.259	0.0453	0.01	0.55
Morphology	Tarsus length	0.256	0.094	0.0275	0.043	0.0796	0.03	1.05
Morphology	Bill length	0.334	0.126	0.0642	1.603	0.1278	0.05	1.07
Morphology	Bill depth	0.287	0.138	0.0408	0.539	0.1015	0.07	2.00
Reproduction	Clutch size	0.150	0.167	0.4058	0.074	2.3101	0.03	0.20
Reproduction	No. hatched	0	0.158	0.0001	0.462	30.4925	0.16	0.23
Reproduction	No. fledged	0	0.146	0.0002	0.505	59.4256	0.18	0.20
Reproduction	× offspring mass	0	0.199	0.0000	2.450	1.4520	0.11	0.97
Reproduction	× offspring wing	0	0.194	0.0000	0.700	1.6807	0.01	0.06
Reproduction	× offspring tarsus length	0	0.109	0.0001	0.192	0.4469	0.04	0.39
Reproduction	Max offspring mass**	0	0.200	0.0000	2.250	1.1515	0.16	1.54
Reproduction	Max offspring wing**	0.167	0.219	0.1431	0.490	0.7118	−0.03	−0.25
Reproduction	Max offspring tarsus^2 length	0	0.121	0.0000	0.191	0.3726	0.09	1.05
Reproduction	Laying date	0.082	0.153	0.0113	5.501	0.1267	−0.22	−7.14
(iv) >1-year-old								
Morphology	Condition (mass/wing)**	0.060	0.204	0.0443	0	0.6944	−0.02	−0.30
Morphology	Condition (mass/tarsus)**	0.017	0.245	0.0129	0.001	0.7672	−0.02	−0.20

(Continued)

Table 1. Continued.

A. Females

Trait type	Trait	h^2	SE h^2	I_A	SE V_A	I_R	i	β_μ
Morphology	Mass	0.140	0.143	0.1072	0.382	0.6571	−0.01	−0.17
Morphology	Wing length	0.505	0.189	0.0274	0.477	0.0269	0.02	1.11
Morphology	Tarsus length	0.518	0.224	0.0422	0.083	0.0393	0.08	2.82
Morphology	Bill length	0.232	0.299	0.0481	4.006	0.1588	0.12	2.62
Morphology	Bill depth	0.177	0.341	0.0240	1.246	0.1116	0.05	1.54
Reproduction	Clutch size	0.056	0.139	0.1627	0.070	2.7689	−0.03	−0.31
Reproduction	No. hatched	0	0.096	0.0015	0.326	37.0394	0.12	0.23
Reproduction	No. fledged	0	0.103	0.0000	0.348	54.6284	0.17	0.24
Reproduction	× offspring mass	0	0.186	0.0000	3.883	2.5163	0.12	1.16
Reproduction	× offspring wing	0.101	0.198	0.2517	1.021	2.2353	0.09	0.73
Reproduction	× offspring tarsus length	0	0.262	0.0000	0.510	0.4951	0.14	2.11
Reproduction	Max offspring mass**	0	0.170	0.0000	3.200	1.9285	0.14	1.73
Reproduction	Max offspring wing**	0.184	0.239	0.2645	0.861	1.1768	0.07	0.64
Reproduction	Max offspring tarsus** length	0.051	0.306	0.0221	0.562	0.4088	0.1	1.61
Reproduction	Laying date	0.135	0.167	0.0218	7.120	0.1398	−0.08	−2.67
Reproduction	Interclutch interval*	0	0.438	0.0001	16.150	3.4868	−0.18	−1.07
Reproduction	Replacement interval*	0		0.0001	5.629	22.1690	0.07	0.15
Reproduction	Postfledging parental care*	0	0.599	0.0000	21.199	6.9464	0.15	0.69
Reproduction	EPP*	0	0.286	0.0000	487.860	0.0000	0.22	0.22
Reproduction	Brood sex ratio*	0	0.199	0.0000	0.0169	0.0000	0.06	0.07
(v) Life history								
Life history	Lifetime mates	0.001	0.040	0.1079	0.027	140.1903	1.65	1.09
Life history	Lifetime nests**	0	0.045	0.0066	0.171	272.2203	1.79	1.06
Life history	Lifetime eggs**	0	0.045	0.0034	2.520	241.8740	1.78	1.03
Life history	Lifetime nestlings**	0	0.046	0.0012	1.860	301.1405	1.87	1.05
Life history	Lifetime fledglings	0	0.052	0.0012	2.020	428.7302	1.92	1.05
Life history	Longevity	0	0.054	0.0138	0.081	127.7601	1.52	1.35
Life history	Lifetime recruits	0.002	0.036	1.5019	0.032	675.9428	2.54	1.00

B. Males

Trait type	Trait	h^2	SE h^2	I_A	SE V_A	IR	i	β_μ
(i) Nestlings								
Morphology	Condition (mass/wing)**	0.045	0.070	0.0560	0	1.1986	0.30	2.85
Morphology	Condition (mass/tarsus)**	0.061	0.102	0.0449	0	0.6853	0.02	0.29
Morphology	Mass	0.099	0.082	0.1549	0.305	1.4102	0.32	2.54
Morphology	Wing length	0	0.092	0.0011	1.840	2.1957	0.09	0.77
Morphology	Tarsus length	0.060	0.357	0.0297	0.735	0.4699	−0.04	−0.62

(Conitnued)

Table 1. Conitnued.

A. Females

Trait type	Trait	h^2	SE h^2	I_A	SE V_A	I_R	i	β_μ
(ii) Juveniles								
Morphology	Condition (mass/wing)**	0.456	0.195	0.1538	0	0.1820	0.10	1.96
Morphology	Condition (mass/tarsus)**	0.109	0.269	0.0440	0	0.3582	0.18	3.08
Morphology	Mass	0.529	0.198	0.1707	0.236	0.1520	0.04	0.66
Morphology	Wing length	0.574	0.113	0.0290	0.312	0.0215	0.08	1.47
Morphology	Tarsus	0.411	0.206	0.0501	0.115	0.0720	−0.04	−0.77
Morphology	Tail length	0	0.402	0.0040	5.424	0.4700	0.07	2.03
(iii) 1-year-old								
Morphology	Condition (mass/wing)**	0.207	0.118	0.0686	0	0.2550	−0.05	−0.93
Morphology	Condition (mass/tarsus)**	0.235	0.132	0.0965	0	0.3148	−0.12	−2.16
Morphology	Mass	0.206	0.115	0.0779	0.180	0.2999	−0.12	−2.18
Morphology	Wing length	0.259	0.099	0.0162	0.298	0.0465	−0.19	−7.58
Morphology	Tarsus length	0.468	0.108	0.0555	0.060	0.0630	−0.05	−1.55
Morphology	Bill length	0.414	0.121	0.0894	1.827	0.1265	−0.08	−1.79
Morphology	Bill depth	0.063	0.094	0.0095	0.403	0.1408	−0.02	−0.55
(iv) >1-year-old								
Morphology	Condition (mass/wing)**	0.255	0.216	0.0853	0	0.2500	0	−0.15
Morphology	Condition (mass/tarsus)**	0.368	0.558	0.1451	0.001	0.2504	−0.11	−2.16
Morphology	Mass	0.380	0.192	0.1342	0.291	0.2193	−0.04	−0.90
Morphology	Wing length	0.452	0.218	0.0232	0.580	0.0280	−0.12	−5.76
Morphology	Tarsus length	0.387	0.305	0.0294	0.109	0.0465	0.08	3.24
Morphology	Bill length	0.097	0.294	0.0200	4.133	0.1865	0.07	1.77
Morphology	Bill depth	0.006	0.292	0.0009	1.366	0.1634	0	−0.23
Reproduction	Postfledging parental care	0.001	0.448	0.0033	6.734	2.7255	−0.12	−0.70
(v) Life history								
Life history	Lifetime mates	0	0.032	0.0460	0.035	181.9181	1.69	0.98
Life history	Lifetime nests**	0	0.032	0.0034	0.165	294.0916	1.80	0.95
Life history	Lifetime eggs**	0	0.029	0.0007	2.060	234.0288	1.81	0.93
Life history	Lifetime nestlings**	0	0.037	0.0018	2.320	350.9066	1.91	1.00
Life history	Lifetime fledglings	0	0.043	0.0003	2.420	475.9291	1.94	1.01
Life history	Longevity	0	0.048	0.0483	0.067	134.6575	1.53	1.34
Life history	Lifetime recruits	0	0.036	0.0203	0.046	723.7876	2.50	0.99

*Combines observations on 1-year-old and >1-year-old females.

**Indicates traits that were excluded because of their high ($r > 0.8$) correlation with at least one other trait.

Table 2. Slopes of linear regressions (± SE) of heritability (h^2) on selection intensity (*i*, absolute value), and of I_A ("evolvability," which is V_A divided by the squared phenotypic mean, multiplied by 100) and the mean-standardized coefficient of residual variance (I_R) on the mean-standardized selection gradient (β_μ, absolute value) for phenotypic traits ($N = 88$, counting features measured at different ages or in different sexes as separate traits) measured on Savannah sparrows on Kent Island, New Brunswick, Canada. Results are broken down by general trait type and sex. Results were similar after removal of highly correlated traits ($N = 60$).

Trait type	Sex	Number of traits	h^2	I_A	I_R
Morphology	Male	25	−0.36 (± 0.47)	−0.003 (± 0.006)	−0.04 (± 0.06)
Morphology	Female	25	−0.50 (± 0.19)	−0.00 (± 0.01)	0.05 (± 0.06)
Reproduction	Female	25	−0.40 (± 0.20)	−0.01 (± 0.02)	−3.37 (± 2.38)
Life history	Male[1]	6	– – –	0.104 (± 0.058)	−367.5 (± 361.7)
Life history	Female	6	0.002 (± 0.001)	0.015 (± 0.173)	−559.0 (± 363.1)
All traits combined	Male[2]	32	−0.14 (± 0.05)	−0.00 (± 0.01)	−0.08 (± 0.08)
All traits combined	Female	56	−0.08 (± 0.03)	−0.00 (± 0.01)	−0.03 (± 0.04)

[1]Slope not calculated because $h^2 = 0$ for all male life history traits.
[2]Includes postfledging parental care by males.

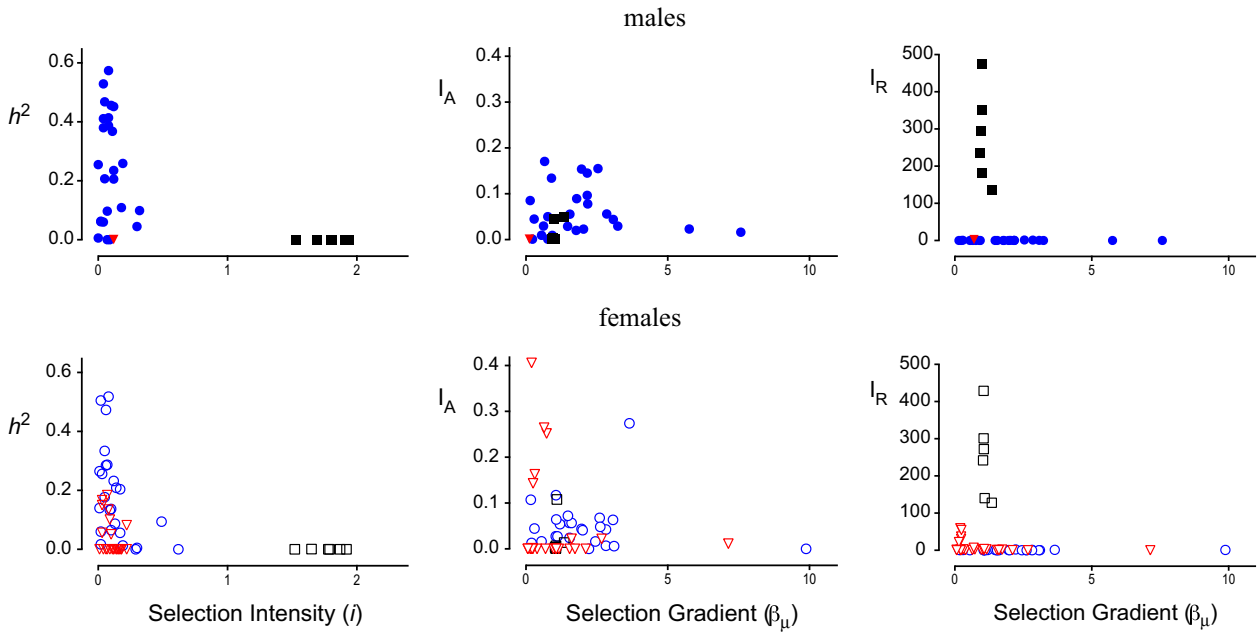

Figure 2. Association of a trait's heritability (h^2) with its variance-standardized selection gradient (β_σ, or selection intensity, *i*), and association of a trait's mean-standardized additive genetic variance (evolvability, I_A) and mean-standardized residual variance (I_R) with its mean-standardized selection gradient (β_μ) in male and female Savannah sparrows. Traits expressed in males = filled symbols; traits expressed in females = open symbols. Morphological traits = circles ($N = 25$ male, 25 female); reproductive traits = triangles ($N = 1$ male, 25 female); life history traits = squares ($N = 6$ male, 6 female) (see Table 1 and Supplementary Tables S1 and S2 for trait descriptions). Note differences in axis scales between I_A and I_R.

among many traits, particularly homologous traits expressed in males and females or at different ages. Nonetheless, in line with previous studies (Kruuk et al. 2000; Merilä and Sheldon 2000; McCleery et al. 2004; Coltman et al. 2005; Teplitsky et al. 2009), we do not include analyses of genetic correlations in this paper or consider the potential role of nonlinear and stabilizing selection in shaping additive genetic variances. Our preliminary attempts at estimating **G** using a subset of traits from this dataset showed that this is an endeavor that is beyond the scope of this paper and our data (see Supplementary Information Online).

In order to maximize precision of estimates and statistical power, we used bivariate (male–female) models for all traits for which we had data for both sexes. Reproductive traits were analyzed only for females (except for duration of postfledging parental care, which is provided by both males and females). Although a female's reproductive behavior is probably influenced by

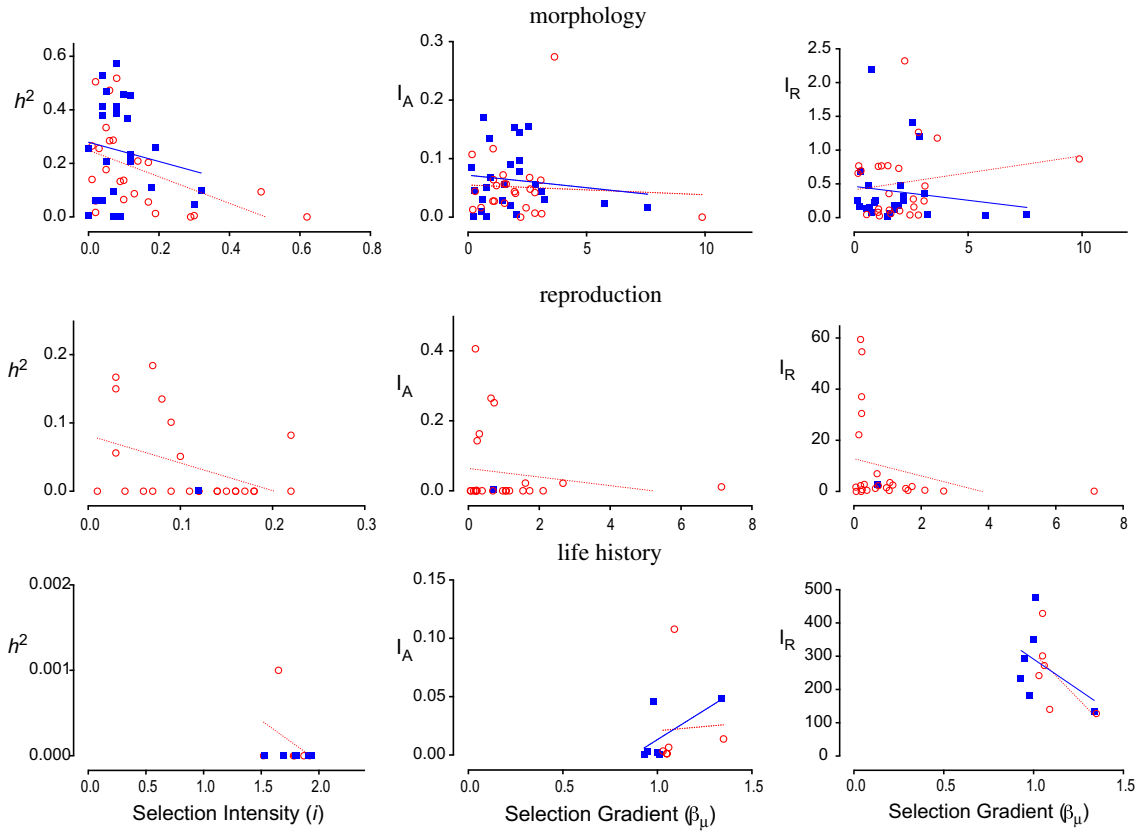

Figure 3. Association of a trait's heritability (h^2) with its variance-standardized selection gradient (β_σ, or selection intensity, i), and association of a trait's mean-standardized additive genetic variance (evolvability, I_A) and mean-standardized residual variance (I_R) with its mean-standardized selection gradient (β_μ) within different trait types (morphology, reproduction, life history). Male traits = filled squares; female traits = open circles (see Fig. 1 and Table 1 and Supplementary Tables S1 and S2). Because points are not strictly independent, regression lines (solid for males, dotted for females) are included simply to indicate trends. Note differences in axis scales between trait types.

the behavior and quality of her mate, females are primarily responsible for choosing nest sites, determining clutch size, incubating, and providing the bulk of feedings to nestlings (Wheelwright and Rising 2008).

To reduce the problem of pseudoreplication due to the inclusion of nonindependent traits, we identified highly correlated traits by examining phenotypic correlation matrices within trait types for females. Correlation coefficients (r) between most pairs of traits were quite low (morphological traits: mean r (SD) = 0.22 (0.20), N = 120 bivariate comparisons; reproductive traits: 0.19 (0.29), N = 61). Correlation coefficients were less than 0.5 for more than 96% of morphological trait pairs and 82% of reproductive trait pairs. On the other hand, certain life history traits (e.g., lifetime eggs and lifetime nestlings) were highly correlated (mean (SD) r for 23 bivariate comparisons = 0.83 (0.11)). To be able to compare our results with previous studies, which did not consider correlations between traits, we included all traits but then repeated our analyses including only traits that had correlation coefficients less than 0.80 with all other traits (see Table 1 and Supplementary Tables S1 and S2).

ESTIMATING QUANTITATIVE GENETICS PARAMETERS

We estimated variance components with a restricted maximum likelihood (REML) procedure and fitted an animal model using the program WOMBAT (Meyer 2006; Wilson et al. 2010). This allowed us to partition phenotypic variance into genetic and environmental components (Wilson et al. 2007, 2010). For additional details on animal models, see Supplementary Information Online.

ESTIMATING STRENGTH OF SELECTION

We estimated the strength of selection on each trait in two ways, using mean-standardized selection gradients (β_μ) and selection intensities, i (i.e., variance-standardized selection gradients, β_σ) (Kruuk et al. 2000; Hereford et al. 2004; Houle et al. 2011). We calculated relative fitness separately for each trait because sample sizes varied depending upon the trait (e.g., N for weights of female

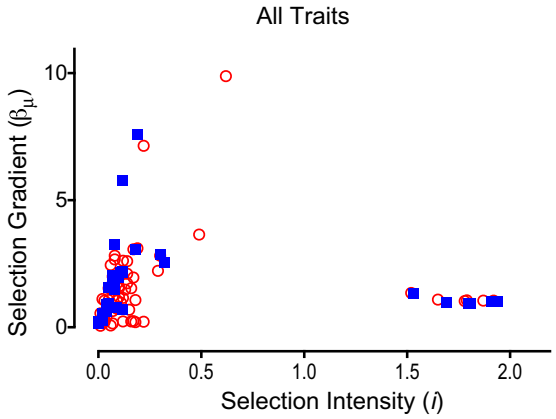

Figure 4. Relationship between mean-standardized selection gradient (β_μ) and variance-standardized selection intensity, *i* (β_σ) for phenotypic traits measured at different ages. Male traits = filled squares; female traits = open circles.

nestlings = 724 whereas N for weights of adult females >1-year-old = 212). An individual's relative fitness was calculated as its lifetime production of recruits divided by the mean lifetime production of recruits for the sample. We chose recruits as a fitness measure rather than fledglings because it is a better predictor of an individual's lifetime genetic contribution to population growth (Brommer et al. 2004). Selection gradients (β) for each trait were calculated as the slope of a linear regression of relative fitness against raw trait values. These were subsequently standardized either by multiplying by trait means (β_μ) or by standard deviations ($\beta_\sigma = i$, selection intensity). Because selection and evolutionary potential should be measured on the same scale, we examined the relationship between h^2 and variance-standardized selection gradients (β_σ), whereas we examined the relationship between I_A and mean-standardized selection gradients (β_μ), as recommended by Hereford et al. (2004) and Matsumara et al. (2012).

Results

Most phenotypic traits measured on Kent Island Savannah sparrows showed extensive variation among individuals. However, heritability tended to be very low, especially for traits such as longevity and other life history features that are influenced by many other traits over the course of a lifetime ("complex traits" sensu Price and Shulter 1991) (Table 1). Because $h^2 = V_A / V_P$, there could still be appreciable additive genetic variance and evolvability, I_A, for traits with low h^2, as long as V_P is high. Nonetheless, we found a median I_A for all traits of 0.024% (mean = 0.050%, SD = 0.072; Table 1, Fig. 1), which is relatively low (see Fig. 5 in Hansen et al. 2011).

Trait types differed in h^2. The mean h^2 (SD) for 50 morphological traits was 0.21 (0.17), compared to 0.036 (0.61) for 26 reproductive traits and only 0.00008 (0.003) for 12 life history traits (Figs. 1 and 3). Traits strongly associated with fitness had the lowest h^2. Mean selection intensity, *i*, for 12 life history traits was 1.80 (0.14), which was substantially higher than *i* for 50 morphological traits (mean = 0.11, SD = 0.12) or for 26 reproductive traits (mean = 0.11, SD = 0.06). I_A did not differ among trait types, unlike I_R, which was much higher for life history traits than for morphological traits (Fig. 1).

Across all traits, h^2 was negatively correlated with the strength of selection as estimated by *i*. In contrast, neither I_A nor I_R was correlated with mean-standardized selection gradients (β_μ) across all traits (Table 2). However, traits with the very highest selection gradients had low I_A and I_R (Figs. 2 and 3). Results were similar when we excluded 28 traits that were highly correlated with at least one other trait (Table 2). Estimates of heritability and evolvability were only weakly correlated (Supplementary Fig. S2; see Hansen et al. 2011).

Although both mean- and variance-standardized selection gradients are designed to measure the strength of selection, they turned out to be only weakly correlated (Fig. 4). In particular, life history traits had high *i* but relatively low β_μ. However, when we excluded life history traits, the two measures of selection were positively correlated but still accounted for only half or less of the variance (females: $R^2 = 0.52$; males: $R^2 = 0.32$) (Fig. 4).

Homologous traits expressed at distinct ages (e.g., wing length as a nestling, juvenile, or adult) differed in h^2 and I_R but not I_A. For all five morphological traits measured at different ages and in each sex, h^2 was substantially lower in nestlings than in independent juveniles, yearlings, or older adults. I_R of morphological traits was higher in nestlings, whereas I_A of morphological traits did not differ among age groups (Fig. 5).

If the mechanism driving the negative relationship between fitness and heritability across all traits is the disproportionate depletion of genetic variation in traits subjected to directional selection, the same negative relationship should hold within trait types for both h^2 and I_A. Although, as described above, we found that h^2 differed *among* trait types and was negatively correlated with the strength of selection, *within* each trait type the relationship was apparent only for female morphological traits and reproductive traits, but not for male morphological traits or for life history traits for either sex (Fig. 2, Table 2). More importantly, within trait types there was no association between I_A and the strength of selection, β_μ. Likewise, I_R showed no relationship with β_μ within trait types (Table 2).

Discussion

In this study of a wild bird population, we have shown that the phenotypic traits that most strongly influence fitness tend to have

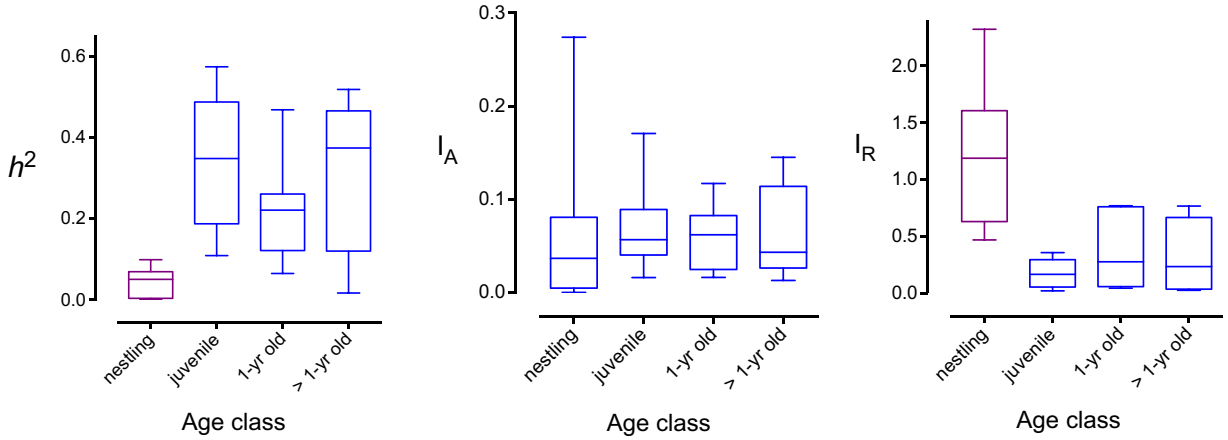

Figure 5. Box and whiskers plots showing heritability (h^2), evolvability (I_A), and mean-standardized residual variance (I_R) for the same morphological traits measured at four different ages ($N = 5$ male and 5 female traits).

low heritability (h^2) but that they do not necessarily have lower evolutionary potential or evolvability (I_A). In that respect, our results agree with the few studies that have estimated quantitative genetic variance components for a range of traits in a natural population (Kruuk et al. 2000; Merilä and Sheldon 2000; McCleery et al. 2004; Coltman et al. 2005; Teplitsky et al. 2009). However, unlike previous studies, our results also indicate that neither h^2 nor I_A is correlated with the strength of selection once trait type (morphology, reproduction, life history) is accounted for. Nor was I_A higher for key life history traits than for traits less closely associated with fitness.

These results are in line with other studies suggesting that the low h^2 commonly observed in traits closely related to fitness is the result of increased residual variance (V_R), rather than reduced additive genetic variance (V_A) (Kruuk et al. 2000; Merilä and Sheldon 2000; McCleery et al. 2004). The negative correlation between h^2 and the strength of selection documented in numerous studies appears to be the result of including in a single analysis such disparate traits as limb length and lifespan (e.g., Gustafsson 1986; Merilä and Sheldon 2000). Thus, our finding that heritability and fitness are negatively correlated across a wide range of disparate traits is unlikely to be due to a Fisherian exhaustion of genetic variation by natural selection (Fisher 1930). The fact that heritability and fitness were not correlated within trait types suggests fundamental differences among traits in how they are shaped by the environment, rather than variation in their association with fitness. Conceivably, selection could erode V_A in fitness traits more than in nonfitness traits, but mutations could at the same time increase V_A more in fitness than in nonfitness traits, with the result that the two processes may cancel each other out, leading to our finding of a lack of relationship between I_A and the strength of selection.

We found that estimates of mean-standardized residual variance (I_R) are much higher in life history traits than in morphological or annual reproductive traits, as expected (Merilä and Sheldon 1999). They are also high in traits expressed at early stages of development (e.g., nestling vs. adult wing length, Fig. 5; see Houle [1992]). This indicates that high residual variance and low h^2 are a function of a trait's accumulated exposure to maternal, nest, and other early environmental effects (Atchley 1984; Houle 1992).

Compared to other studies of vertebrates in the wild, our estimates of h^2 are lower for most traits. This could reflect fundamental differences in the quantitative genetics of different populations and species, but it may also be partly due to errors in our pedigree introduced by undetected extra-pair paternity (Freeman-Gallant et al. 2005), or to the fact that we employed animal models rather than traditional parent-offspring regressions (Merilä and Sheldon 1999), which are prone to inflating estimates of h^2 (Merilä and Sheldon 2001; Kruuk and Hadfield 2007). Moreover, environmental variance is likely to be greater in truly natural populations such as Savannah sparrows on Kent Island than in populations provided with homogeneous conditions for reproduction (e.g., uniform artificial nest boxes) (McCleery et al. 2004) or subjected to culling or other human management (Coltman et al. 2005). Higher environmental variance would result in lower estimates of heritability (Falconer and Mackay 1996).

In conclusion, our results indicate that certain types of traits are likely to have low heritability because of an increase in residual variance due to the age at which they are expressed and the many other nonadditive genetic factors that influence it, rather than variation in the strength of (directional) selection acting upon them. This study emphasizes that when investigating how natural selection shapes additive genetic variance and evolutionary potential, distinguishing between trait types and taking into account trait ontogeny and complexity, as envisioned by Price and Schluter (1991), is essential.

ACKNOWLEDGMENTS

We are grateful to our reviewers for their insightful comments and for directing us to several key references. This study was made possible thanks to support over 24 years to NTW from the U.S. National Science Foundation (NSF) and Bowdoin College, as well as the dedication of numerous field assistants. We particularly express our appreciation to Beat and Elisabeth Messerli for valuable discussions, Corey Freeman-Gallant for molecular paternity assignments, and NSF for OPUS award no. 0816132 (to NTW). EP and LFK are funded by the Swiss National Science Foundation (SNSF) (grant 31003A–116794). The authors have no conflict of interest. The data reported in this paper are archived in the Savannah Sparrow database on the website of the Bowdoin Scientific Station and Dryad. This represents contribution no. 248 from the Bowdoin Scientific Station.

DATA ARCHIVING

The doi for our data is 10.5061/dryad.dv0qt.2.

LITERATURE CITED

Atchley, W. R. 1984. Ontogeny, timing of development, and genetic variance-covariance structure. Am. Nat. 123:519–540.

Blows, M. W. 2006. A tale of two matrices: multivariate approaches in evolutionary biology. Evol. Biol. 20:1–8.

Brommer, J. E., L. Gustafsson, H. Pietianinen, and J. Merilä. 2004. Single-generation estimates of individual fitness as proxies for long-term genetic contribution. Am. Nat. 163:505–517.

Brommer, J. E., J. Merilä, B. C. Sheldon, and L. Gustafsson. 2005. Natural selection and genetic variation for reproductive reaction norms in a wild bird population. Evolution 59:1362–1371.

Charmantier, A., and D. Reale. 2005. How do misassigned paternities affect the estimation of heritability in the wild? Mol. Ecol. 14:2839–2850.

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

Falconer, D. S., and T. F. C. Mackay. 1996. Introduction to quantitative genetics. Longman, Essex, England.

Fisher, R. A. 1930. The genetical theory of natural selection. Clarendon Press, Oxford, U.K.

Freeman-Gallant, C. R., M. Meguerdichian, N. T. Wheelwright, and S. V. Sollecito. 2003. Social pairing and female mating fidelity predicted by restriction fragment length polymorphism similarity at the major histocompatibility complex in a songbird. Mol. Ecol. 12:3077–3083.

Freeman-Gallant, C. R., N. T. Wheelwright, K. E. Meiklejohn, S. L. States, and S. V. Sollecito. 2005. Little effect of extrapair paternity on the opportunity for sexual selection in Savannah sparrows (*Passerculus sandwichensis*). Evolution 59:422–430.

Freeman-Gallant, C. R., N. T. Wheelwright, K. E. Meiklejohn, and S. V. Sollecito. 2006. Genetic similarity, extrapair paternity, and offspring quality in Savannah sparrows (*Passerculus sandwichensis*). Behav. Ecol. 17:952–958.

Garcia-Gonzalez, F., L. W. Simmons, J. L. Tomkins, J. S. Kotiaho, and J. P. Evans. 2012. Comparing evolvabilities: common errors surrounding the calculation and use of coefficients of additive genetic variation. Evolution 66:2341–2349.

Gustafsson, L. 1986. Lifetime reproductive success and heritability: empirical support for Fisher's Fundamental Theorem. Am. Nat. 128:761–764.

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

Hereford, J., T. F. Hansen, and D. Houle. 2004. Comparing strengths of directional selection: how strong is strong? Evolution 58:2133–2143.

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

———. 1998. How should we explain variation in the genetic variance of traits? Genetica 102--103:241–253.

Houle, D., B. Morikawa, and M. Lynch. 1996. Comparing mutational variabilities. Genetics 143:1467–1483.

Houle, D., C. Pélabon, G.P. Wagner, and T.F. Hansen. 2011. Measurement and meaning in biology. Quart. Rev. Biol. 86:3–34.

Hunt, J., M. W. Blows, F. Zajitschek, M. D. Jennions, and R. Brooks. 2007. Reconciling strong stabilizing selection with the maintenance of genetic variation in a natural population of black field crickets (*Teleogryllus commodus*). Genetics 177:875–880.

Jensen, H., B. E. Saether, T. H. Ringsby, J. Tufto, S. C. Griffith, and H. Ellegren. 2003. Sexual variation in heritability and genetic correlations of morphological traits in house sparrow (*Passer domesticus*). J. Evol. Biol. 16:1296–1307.

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

Kimura, M. 1958. On the change of population fitness by natural selection. Heredity 12:145–167.

Kingsolver, J. G., H. E. Hoekstra, J. M. Hoekstra, D. Berrigan, S. N. Vignieri, C. E. Hill, A. Hoan, P. Gilbert, and P. Beerli. 2001. The strength of phenotypic selection in natural populations. Am. Nat. 157:245–261.

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

Kruuk, L. E. B., and J. D. Hadfield. 2007. How to separate genetic and environmental causes of similarity between relatives. J. Evol. Biol. 20:1890–1903.

Kruuk, L. E. B., T. H. Clutton-Brock, J. Slate, J. Pemberton, S. Brotherstone, and F. E. Guiness. 2000. Heritability of fitness in a wild mammal population. Proc. Natl. Acad. Sci. 97:698–703.

Lande, R., and S. J. Arnold. 1983. The measurement of selection on correlated characters. Evolution 37:1210–1226.

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

Matsumara, S., R. Arlinghaus, and U. Dieckmann. 2012. Standardizing selection strengths to study selection in the wild: a critical comparison and suggestions for the future. BioScience 62:1039–1054.

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.

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

———. 2000. Lifetime reproductive success and heritability in nature. Am. Nat. 155:301–310.

———. 2001. Avian quantitative genetics. Curr. Ornithol. 16:179–255.

Meyer, K. 2006. WOMBAT—digging deep for quantitative genetic analyses by restricted maximum likelihood. Pp. 27–14 *in* Proceedings of the 8th World Congress in Genetic Applied Livestock Production, Belo Horizonte, Brazil.

Mousseau, T. A., and D. A. Roff. 1987. Natural selection and the heritability of fitness components. Heredity 59:181–197.

Postma, E. 2014. Four decades of estimating heritabilities in wild vertebrate populations: improved methods, more data, better estimates? *In* A. Charmantier, D. Garant, and L. E. B. Kruuk, eds. Quantitative genetics in the wild. Oxford Univ. Press, Oxford, U.K.

Postma, E., N. Spyrou, L. A. Rollins, and R. C. Brooks. 2011. Sex-dependent selection differentially shapes genetic variation on and off the guppy Y chromosome. Evolution 65:2145–2156.

Price, T., and D. Schluter. 1991. On the low heritability of life-history traits. Evolution 45:853–861.

Steven, J. C., L. F. Delph, E. D. Brodie, and I. I. I. C. Fenster. 2007. Sexual dimorphism in the quantitative-genetic architecture of floral, leaf and allocation traits in *Silene latifolia*. Evolution 61:42–57.

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

Teplitsky, C., M. Tarka, A. P. Møller, S. Nakagawa, J. Balbontin, T. A. Burke, C. Doutrelant, A. Gregoire, B. Hansson, D. Hasselquist, et al. 2014. Assessing multivariate constraints to evolution across ten long-term avian studies. PLoS One 9. doi:10.1371/journal.pone.0090444.

Wheelwright, N. T., and R. A. Mauck. 1998. Philopatry, natal dispersal, and inbreeding avoidance in an island population of Savannah sparrows. Ecology 79:755–767.

Wheelwright, N. T., and J. D. Rising. 2008. Savannah Sparrow (*Passerculus sandwichensis*). The Academy of Natural Sciences and Cornell Laboratory of Ornithology, Ithaca, New York.

Wheelwright, N. T., C. R. Freeman-Gallant, and R. A. Mauck. 2006. Asymmetrical incest avoidance in the choice of social and genetic mates. Anim. Behav. 71:631–639.

Wilson, A. J., J. M. Pemberton, J. G. Pilkington, T. H. Clutton-Brock, D. W. Coltman, and L. E. B. Kruuk. 2007. Quantitative genetics of growth and cryptic evolution of body size in an island population. Evol. Ecol. 21:337–356.

Wilson, A. J., D. Reale, M. N. Clements, M. M. Morrissey, E. Postma, C. A. Walling, L. E. B. Kruuk, and D. H. Nussey. 2010. An ecologist's guide to the animal model. J. Anim. Ecol. 79:13–26.

Associate Editor: A. Charmantier

Supporting Information

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

Figure S1. Comparison of h^2 estimates based on animal models versus traditional single parent-offspring regressions in Savannah sparrows.

Figure S2. Plot of heritability against evolvability in Savannah sparrows.

Table S1. Description and methods of measuring phenotypic traits of Savannah sparrows on Kent Island, New Brunswick, Canada.

Table S2. Descriptive statistics for phenotypic traits measured on Savannah sparrows on Kent Island, New Brunswick, Canada (see Supplementary Table S1 for units).