Evol Biol (2011) 38:258–277
DOI 10.1007/s11692-011-9127-6

RESEARCH ARTICLE

Heritability is not Evolvability

Thomas F. Hansen · Christophe Pélabon · David Houle

Received: 9 March 2011 / Accepted: 2 June 2011 / Published online: 21 June 2011

Abstract Short-term evolutionary potential depends on the additive genetic variance in the population. The additive variance is often measured as heritability, the fraction of the total phenotypic variance that is additive. Heritability is thus a common measure of evolutionary potential. An alternative is to measure evolutionary potential as expected proportional change under a unit strength of selection. This yields the mean-scaled additive variance as a measure of evolvability. Houle in Genetics 130:195–204, (1992) showed that these two ways of scaling additive variance are often inconsistent and can lead to different conclusions as to what traits are more evolvable. Here, we explore this relation in more detail through a literature review, and through theoretical arguments. We show that the correlation between heritability and evolvability is essentially zero, and we argue that this is likely due to inherent positive correlations between the additive variance and other components of phenotypic variance. This means that heritabilities are unsuitable as measures of evolutionary potential in natural populations. More generally we argue that scaling always involves non-trivial assumptions, and that a lack of awareness of these assumptions constitutes a systemic error in the field of evolutionary biology.

Keywords Evolvability · Heritability · Genetic variance · Quantitative genetics · Measurement theory · Scaling

Introduction

A focal question for evolutionary quantitative genetics is whether, or at what time scales, variation in evolutionary potential is important in predicting evolutionary change. One view is that capacities to evolve are typically high enough to allow populations to closely track adaptive optima even on time scales as short as tens or hundreds of generations. In this case we do not expect variation in evolutionary potential across traits and populations to have much explanatory value, and macroevolutionary patterns can be understood mainly in terms of selection dynamics (e.g. Schluter 2000; Arnold et al. 2001; Estes and Arnold 2007; Futuyma 2010). The opposite view is that various forms of variational constraints can influence trait dynamics over longer macroevolutionary time scales (e.g. Hansen and Houle 2004; Blows and Hoffmann 2005; Kirkpatrick 2009; Walsh and Blows 2009). In this case, we predict that evolutionary changes in traits or trait combinations with higher evolvability may be larger and more frequent. Understanding the dynamics and determinants of evolvability then becomes essential for understanding macroevolutionary dynamics and patterns.

To make empirical progress on such questions, it is instrumental to develop operational measures of

Electronic supplementary material The online version of this article (doi:10.1007/s11692-011-9127-6) contains supplementary material, which is available to authorized users.

T. F. Hansen (✉)
Department of Biology, Centre for Ecological and Evolutionary Synthesis, University of Oslo, PB1066, Blindern, 0316 Oslo, Norway
e-mail: Thomas.Hansen@bio.uio.no

C. Pélabon
Department of Biology, Centre for Conservation Biology, Norwegian University of Science and Technology, 7491 Trondheim, Norway

D. Houle
Department of Biological Science, Florida State University, Tallahassee, FL 32306, USA

evolvability. Perhaps symptomatic of a general lack of attention to measurement issues in biology (Houle et al. 2011), the measurement of evolvability has received little attention. There is a rich literature on definitions, dynamics and determinants of evolvability (see e.g. Wagner and Altenberg 1996; Schlichting and Murren 2004; Hansen 2006; Willi et al. 2006; Hendrikse et al. 2007; Polly 2008; Brookfield 2009 for review), but there has been little discussion of how to quantify the concept. The authors of this paper have struggled for many years in various combinations to develop a measurement theory for evolvability and related concepts (e.g. Houle 1992, 1998; Hansen et al. 2003a, b; Hereford et al. 2004; Carter et al. 2005; Hansen and Houle 2008; Houle et al. 2011). We have focused on short-term measures of evolvability that are meaningful within the context of "additive" evolutionary quantitative genetics (e.g. Lande 1976, 1979; Wagner 1989). This is not because we think short-term evolvability can be naively extrapolated, but because the study of evolvability dynamics on longer time scales has to build on a solid understanding of evolvability on short time scales. Indeed, there is an increasing understanding of how properties of the genotype-phenotype map such as pleiotropy and epistasis may influence the evolution of additive effects and thus evolvability (reviewed in Hansen 2006). We note that there has been a recent surge of interest in theory-based measures of evolvability and related concepts such as modularity and integration (Blows et al. 2004; Mitteroecker and Bookstein 2007, 2009; Wagner et al. 2007; Hansen and Houle 2008; Stinchcombe et al. 2008; Agrawal and Stinchcombe 2009; Gomulkiewicz and Houle 2009; Hallgrimsson et al. 2009; Kirkpatrick 2009; Marroig et al. 2009; Pavlicev et al. 2009; Walsh and Blows 2009).

In this paper we will consider a simple but fundamental issue of measurement, namely standardization. If evolutionary potential is to be compared across traits and species, it needs to be measured in a way that means roughly the same thing in the different situations. If we define evolvability as the expected evolutionary response to selection per strength of selection, we see that the proper standardization of evolvability follows from how we measure and standardize the evolutionary response and the strength of selection (Hansen and Houle 2008). There are, however, several ways of doing this, and thus several alternative ways of measuring evolvability. At the core of evolutionary quantitative genetics sits the simple "Lande equation", which in the univariate case says that the expected change in the trait mean per generation is $R = V_A\beta$, where V_A is the additive genetic variance in the trait and β is the selection gradient, the change in relative fitness per change in the trait. In the context of this equation, the evolvability is measured as $e = R/\beta = V_A$. The additive variance, however, has units equal to the square of the trait units and need be standardized for comparison across traits. This has commonly been done by dividing the additive variance with the total phenotypic variance. This yields the heritability, h^2, which is also the standard measure of evolutionary potential in animal and plant breeding (Falconer and MacKay 1996), in which it derives its relevance from the breeder's equation, $R = h^2S$, where S is the selection differential, the covariance between the trait and relative fitness. Thus, heritability is a variance-standardized measure of evolvability as $e_\sigma = R/S = h^2$.

The heritability has, however, proven to be a problematic measure of evolutionary potential (Fisher 1951; Burton 1952; Houle 1992; Hansen et al. 2003b; Wilson 2008), and Houle (1992) proposed to use mean scaling rather than variance scaling for comparing evolvabilities. Following Hansen and Houle (2008), the mean-scaled evolvability is $e_\mu = V_A/m^2$, where m is the trait mean before selection. Houle (1992) compared variance-scaled and mean-scaled additive variances, and showed that they may lead to very different conclusions. For example, while heritabilities tend to be lower for life-history traits than for morphological traits (Roff and Mousseau 1987; Mousseau and Roff 1987), the situation is the reverse for mean-scaled additive variances. Our interpretation of this is that life-history traits indeed tend to have high levels of additive variance, but even higher levels of total variance. Thus, the heritability fails as a measure of evolvability in this comparison, because the standardization with phenotypic variance acts as a rubber scale that get stretched when we measure something large. This is not the only case in which we expect different conclusions from mean and variance scaling. As shown by a plot of Houle's (1992) data in Fig. 1, the correlation between the two measures is almost exactly zero.

In this paper, we further explore this surprising and dramatic effect of scaling additive variance. We capitalize on the increased reporting and awareness of mean-scaled evolvabilities following the publication of Houle (1992). By surveying all issues of Evolution and Journal of Evolutionary Biology from 1992 to 2009, we collected 1,465 estimates of mean-scaled and variance-scaled genetic variances. These data confirm a near zero correlation between heritabilities and mean-scaled additive variances both in general and within specific studies and trait categories. We discuss the reasons for and the implications of this finding. We conclude that the choice of scale generally can, and often will, have dramatic effects on the conclusions drawn. This is underappreciated in evolutionary biology. The choice of scale entails strong biological assumptions, and should be considered an integral part of model building.

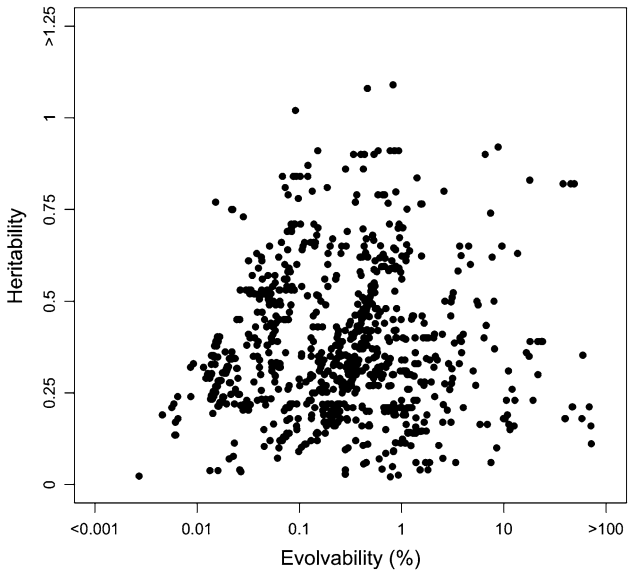

Fig. 1 Plot of heritability against evolvability (=mean-scaled additive genetic variance) for Houle's (1992) data (excluding non-positive estimates). The correlation is -0.03 ± 0.04 for positive values (and 0.01 ± 0.04 if non-positive estimates are included). The correlation between heritability and additive variance on $\log_{10}$-scale, as shown, is 0.06 ± 0.03

Theoretical Background

Hansen and Houle (2008) have developed a theoretical framework for the measurement of multivariate evolvability and associated parameters. We will follow this framework, but only consider the univariate case. The starting point is to measure evolvability as the response to selection per strength of selection. This is only operational in the context of a specific model of selection and its evolutionary response. We use the "Lande equation", which, as explained in the introduction, describes the expected change in the trait mean from one generation to the next under linear directional selection. In this framework the evolvability is

$$e = R/\beta, \tag{1}$$

where R is the change in trait mean from generation to generation, and β is the selection gradient. Since the response is measured in units of trait change per generation, and the selection gradient has units equal to the inverse of the units of the trait per generation, the raw evolvability has units of the trait squared. From the Lande equation we can then infer that $e = V_A$, where V_A is the additive genetic variance. The two main alternatives for standardizing this are

$$e_\sigma = V_A/V_P = h^2, \tag{2}$$

$$e_\mu = V_A/m^2 = I_A, \tag{3}$$

where V_P is the total phenotypic variance, m is the mean of the trait, and I_A is the mean-standardized additive variance. Note that we will use the notation e_σ and e_μ to denote standardized evolvabilities defined as responses per strength of selection, while we will use h^2 and I_A to denote standardized additive variances, which can be meaningful in other ways than as measures of evolvability. Note also that I_A is equal to the square of the coefficient of additive genetic variance, CV_A, which was the measure of evolvability emphasized by Houle (1992). While CV_A may be a more familiar measure of variation, I_A, or more precisely e_μ, is preferable as a measure of evolvability, since its numerical value has a more direct interpretation as the expected percent change in a trait under a unit strength of selection (Hansen et al. 2003b; see also Houle 1992; Sgro and Hoffmann 1998).

The scaling of evolvability must also be seen in the context of scaling the other components of the evolutionary model. The mean-scaled Lande equation is

$$R/m = \left(V_A/m^2\right)(\beta m) = e_\mu \beta_\mu, \tag{4}$$

where $\beta_\mu = \beta m$ is the mean-scaled selection gradient (Hansen et al. 2003b; Hereford et al. 2004). The mean-scaled selection gradient measures the proportional change in fitness with a proportional change in the trait, and is technically an elasticity (van Tienderen 2000). Note that $\beta_\mu = 1$ means that a 1% change in the trait yields a 1% change in fitness. This would be the strength of selection on fitness itself as a trait, which is an invariant in evolutionary theory. This also yields an interpretation of e_μ as the expected percent change in the trait mean if the trait was subject to unit selection ($\beta_\mu = 1$).

The variance-scaled Lande equation is

$$R/\sqrt{V_P} = (V_A/V_P)(\beta\sqrt{V_P}) = e_\sigma \beta_\sigma, \tag{5}$$

where $\beta_\sigma = \beta\sqrt{V_P}$ is the variance-scaled selection gradient. It measures the proportional change in fitness with a change of one standard deviation in the trait. Note that $\beta_\sigma = i$, where i is the selection intensity. We can obtain the same result by starting from the breeder's equation. The variance-standardized breeder's equation is

$$R/\sqrt{V_P} = h^2(S/\sqrt{V_P}) = h^2 i, \tag{6}$$

which, since $\beta = S/V_P$, is mathematically equivalent to the above with $e_\sigma = h^2$ and $\beta_\sigma = i$.

At this point the reader may wonder why we bother to spell out several mathematically equivalent formulations in a variety of forms and symbols. The reasons for this are as follows: First, it shows how different formulations and symbols used in the literature relate to each other. Second, the different formulations entail different implicit assumptions about what entities are related to each other

and can be grouped together as quasi-autonomous conceptual units. For example, the breeder's equation entails the assumption that the selection differential and the heritability are quasi-independent units in the sense that we expect the response to selection to be proportional to their values taken in isolation. As we will argue in the discussion, this assumption is problematic since both the selection differential and the heritability depend on the environmental variance, which causes a negative correlation between them. This implies that a large selection differential or intensity, or a high heritability will not necessarily lead to a large evolutionary response. Third, the different formulations entail different assumptions about what combinations of mathematical symbols map to verbal concepts such as evolutionary potential and strength of selection. While the selection gradient is a measure of the steepness of the fitness landscape, the mean-scaled gradient β_μ measures this in units of the trait mean, while the variance-scaled β_σ measures it in units of the trait standard deviation. The former assumes that doubling of a trait will have a comparable effect on fitness regardless of the size of the trait, while the latter assumes that a standard deviation change of the trait will have a comparable effect on fitness regardless of the level of variation in the trait. Clearly, these are different biological assumptions. How we cut our conceptual cake can have profound consequences for our perspective on the process we seek to model, and the empirical consequences of these choices can be dramatic.

Methods

We surveyed all issues of Evolution and Journal of Evolutionary Biology from 1992 to 2009 for studies that report sufficient statistics to calculate mean-scaled additive genetic variances and heritabilities for traits on comparable scales. As our main goal was to study the relation between these measures as used in the field, we tried to include everthing the authors presented as measures of heritability or additive variance and we did not try to judge or exclude data based on the quality of the methods or data except when there were obvious major errors of reporting. Even with this relaxed standard, the majority of quantitative genetic studies in these journals could not be used either because they did not report necessary statistics or due to ambiguities in the methods with regard to scale or scale type of the measurements. For similar reasons, many individual estimates were excluded from studies that we did use.

A fair comparison of the two measures puts constraints on the scale type of the data we can use. Scaling with the mean is meaningful only for data on a ratio or log-interval scale, and scaling with the variance is meaningful only for data on a ratio, interval or difference scale (Houle et al. 2011). For this reason we included only traits on a ratio scale and also traits on a log-interval scale when it was clear that these were treated as ratio scales within the study. A log-interval scale is a scale that allows changes of both units and dimension (through power transformations), and may include traits such as size, which can be measured either linearly, as area or as volume. The variance of an area and a volume are not directly comparable, but as long as we are comparing heritabilities and evolvabilities on data on the same scale, we can regard this as a difference similar to the difference between qualitatively different traits, which we do compare.

Traits are also often log transformed, which maps ratio and log-interval scales to difference and interval scales. Mean-scaled evolvabilities are thus meaningless on a log scale. We can, however, use such data since the variance of the log-transformed data is to a first approximation equal to the mean-scaled variance on the original scale. If $I[x] = Var[x]/E[x]^2 \ll 1$, then

$$Var[Log[x]] \approx I[x]. \quad (7)$$

Therefore we can use the additive variance of the log of a trait as an estimate of I_A for the trait on the original scale. Indeed, the additive variance on log scale may be regarded as an alternative measure of evolvability on a scale of proportional change. Note, however, that this can only be compared to a heritability on the original scale, and not to a heritability computed on the log scale. Some studies could not be used for this reason.

We included studies using many different designs. Additive genetic variance components are usually computed from the resemblance of relatives in some breeding design. Most studies used half-sib designs, but we also included studies based on parent-offspring regressions, diallels, and more complex designs with different types of relatives. After about year 2000, quantitative genetic mixed models (Lynch and Walsh 1998) start to be used, and are quite common in later studies. We also included a few realized heritabilities from artificial-selection experiments. When proper additive variances and narrow-sense heritabilities were not available, we included estimates from full-sib studies and also broad-sense heritabilities and total mean-scaled genetic variances. We did this because our main focus was to compare mean and variance scaling, which may have similar effects on all genetic variance components (but see discussion). We include a supplementary file that compares results for some specific designs. When a study reported estimates from alternative methods for the same trait, we used the estimate we deemed to be the most reliable. When estimates for the same trait were available for many replicate populations, we used an average of these. We did, however, include

estimates from males and females and from any clearly distinct populations as separate observations.

We classified each estimate into different trait categories. The category of size includes all morphological measurements that could be expected to scale monotonically with size for purely mechanical reasons. This was further divided into linear, area, volume, weight, and count measures. Morphological measurements that involved ratios, angles or other non-monotonic functions of size where classified as shape traits, and measurements on colors or other body patterns were classified as pattern traits. Life-history traits include measurements of fertility, timing of development, survival, and other fitness components. Measurements of size change per developmental time were classified as growth traits. A variety of measurements of physiological and biochemical processes were classified as physiological traits and a variety of behavioral measures were classified as behavioral traits. Fluctuating asymmetry and related measures were classified as developmental-stability traits.

Some studies report heritabilities but not genetic variance components. We used these studies if some estimate of phenotypic variance was available that allowed us to compute additive genetic variance from the heritability. We could often compute phenotypic variances from reported standard deviations, and in a few cases from reported standard errors of the mean when sample sizes were available. We did not use variances based on estimates that were reported with only a single significant digit.

Since we were interested in how well heritabilities and evolvabilities can predict each other, we focused on the squared Pearson correlation coefficient, R^2, between mean- and variance-scaled estimates. This tells us how much variance in one variable is explained by the other in a linear regression forced through zero. Because biological meaning resides on the original scale, we compute our statistics on the original scales and not on the $\log_{10}$-scales used for plotting (correlations tended to be slightly larger with evolvabilities on the $\log_{10}$ scale). Unless otherwise indicated we excluded zero and negative evolvabilities when we computed the correlations. Hence, our R^2 pertains to how well positive heritabilities and evolvabilities can predict each other. We assessed sampling uncertainty by nonparametric bootstrapping (10,000 replicates) on the level of individual data points and report these as estimates $\pm$ standard error.

Results

Our survey yielded 1,465 estimates of mean-scaled and variance-scaled genetic variances from a total of 157 studies. The studies are listed in Appendix, and the data we used are given in supplementary appendix S1. A list of quantitative genetic studies that we considered but did not include is given in supplementary appendix S2. Figure 2 confirms the finding from Houle's (1992) data that the correlation between mean- and variance-scaled genetic variances is very small with $R^2 = 1.3 \pm 0.6\%$, as compared to $R^2 = 0.0 \pm 0.003\%$ in Houle's (1992) data. Despite the lack of overall correlation, both Figs. 1 and 2 indicate a weak positive relationship when heritabilities are small, and this is also seen in the slightly higher correlation on log scale where the effects of small estimates are enhanced. This is not unexpected, since the correlation between additive variance and total variance is presumably lower when the additive variance is a smaller component of the latter, and a zero or negative heritability implies a zero or negative evolvability.

It is possible that the absence of an overall correlation is due to the heterogeneity of traits and methods in the different studies, and that stronger relationship may exist in narrower context of specific trait categories or across traits within specific populations. In Figs. 3 and 4 we investigate this possibility. In Fig. 3 we plot the relationships within different trait categories. The most striking observation is that the observed correlation is almost exactly zero for linear size measures ($R^2 = 0.1 \pm 0.3\%$). This is probably the trait category most likely to be comparable across organisms, and is also based on a large sample size. The results are similar for all size traits combined ($R^2 < 1\%$), and for weight-based size measures ($R^2 = 3 \pm 4\%$). There

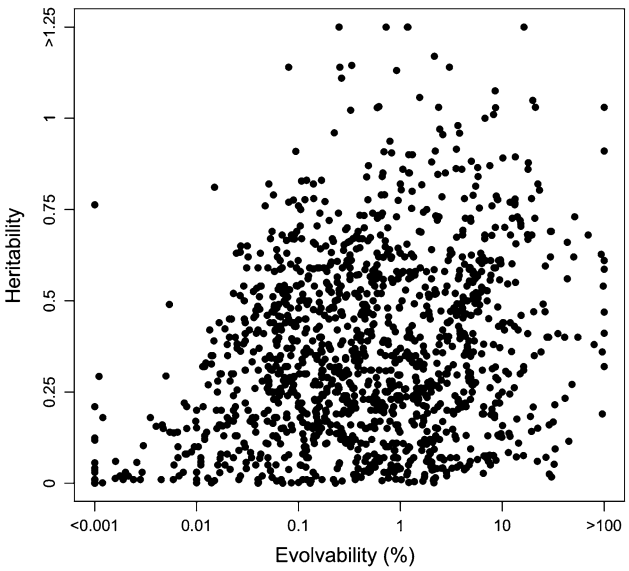

Fig. 2 Plot of heritability against evolvability (=mean-scaled additive genetic variance) for the entire data set (excluding non-positive estimates). The correlation is 0.11 ± 0.02 for positive values (and 0.13 ± 0.02 if non-positive estimates are included). The correlation between heritability and evolvability on $\log_{10}$-scale, as shown, is 0.26 ± 0.03

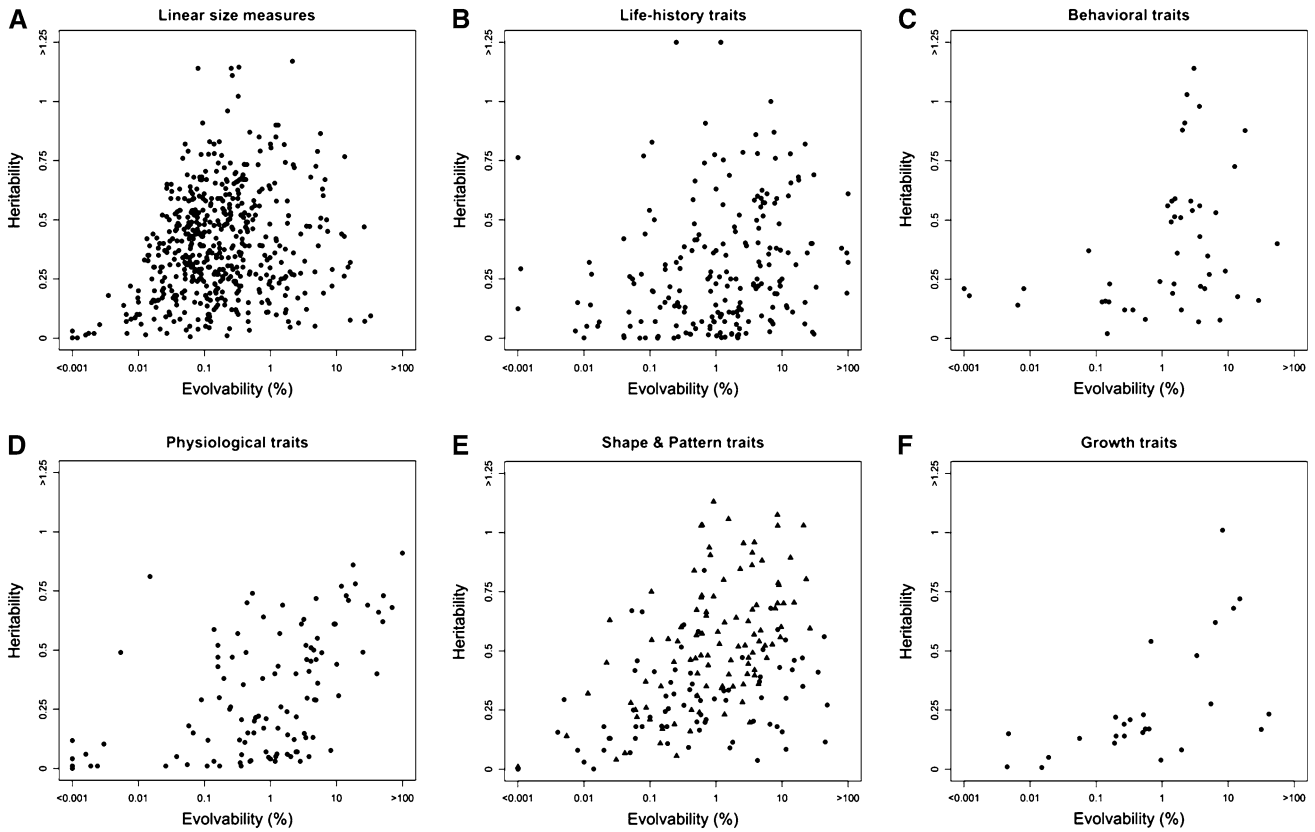

Fig. 3 Plots of heritability against evolvability for specific trait categories: **a** Linear size measurements (correlation = −0.03 ± 0.05). **b** Life-history traits (correlation = 0.09 ± 0.05). **c** Behavioral traits (correlation = 0.06 ± 0.12). **d** Physiological traits (correlation = 0.47 ± 0.05). **e** Shape (*circles*) and Pattern (*triangles*) traits (correlation = 0.13 ± 0.08). **f** Growth traits (correlation = 0.25 ± 0.22). Correlations are for positive estimates only

are higher correlations for size on other scales (R^2 = 36 ± 15%, and 10 ± 8% for area and count, respectively), but these involve much smaller sample sizes. For life-history and behavioral traits, as well as morphological shapes and patterns, the result is the same with R^2 around 1%. But for physiological and growth traits there is a signal. For physiological traits, with a decent sample size, the R^2 is 22 ± 5%, although this is still much too small to allow prediction with reasonable accuracy. The only category for which we found a strong correlation was developmental stability traits, i.e. fluctuating asymmetry, with R^2 = 84 ± 21%. This is due to many heritabilities in the vicinity of zero.

In Fig. 4 we plot the relationship between heritability and mean-scaled genetic variances for some selected studies and organisms. In Fig. 4a we show the results for Cheverud's (1996) study of cranial measurements in two species of Tamarins. This is probably the one study in our dataset best suited to generate a relationship. It is based on large samples of traits and individuals, traits are strictly comparable, and the methods of measurement and analysis are exemplary. Although there is a clear relationship with R^2 = 38 ± 8%, a prediction of evolvability from heritability or vice versa would be crude even in this case. A less strong relation was found in our own two studies of floral variation in *Dalechampia* shown in Fig. 4b. Here there is a moderately positive correlation (R^2 = 34 ± 8%), but much of this is driven by negative estimates of additive genetic variance for fluctuating asymmetries (R^2 = 14 ± 9% for positive estimates only). A different result is shown in Fig. 4c from several studies of life-history and size-related traits in big-horn and domestic sheep. These are methodologically strong studies based on sophisticated mixed-model analyses of pedigrees, but the traits are more different, and the correlation between heritability and evolvability is in fact negative. A similar weak negative correlation was found for linear size measures in *Drosophila* as illustrated in Fig. 4d.

In Table 1 we report median heritabilities and evolvabilities for the various trait categories in our study; there is no tendency for groups of traits with high evolvabilities to have high heritabilities. We confirm Houle's (1992) observation that life-history traits tend to have high evolvabilities and low heritabilities while the situation is reversed for morphological traits (Size and Shape in our classification). The median evolvability for life-history

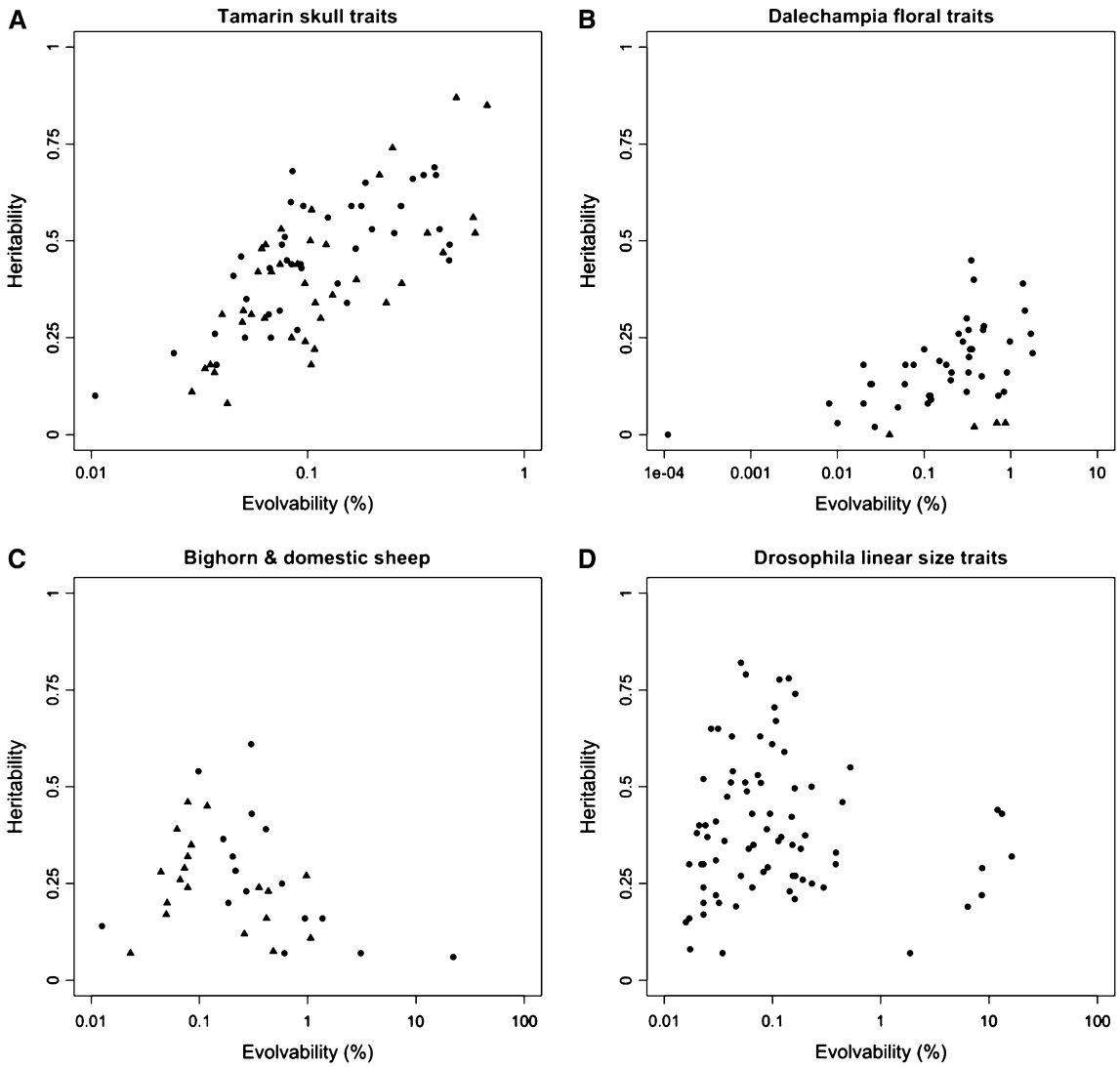

Fig. 4 Plots of heritability against evolvability for specific studies and organisms: **a** Linear skull measurements from two species of tamarins (*Saguinus oedipus* (*circles*) and *S. fusicollis* (*triangles*)) from Cheverud (1996). Here the correlation is 0.62 ± 0.07. **b** Floral measurements from a greenhouse population of *Dalechampia scandens*. The *circles* are morphological traits from Hansen et al. (2003b), and the *triangles* are fluctuating asymmetries from Pélabon et al. (2004), six negative additive genetic variances from the latter study were not included in the plot. The correlation is 0.43 ± 0.12 (and 0.58 ± 0.08 if the negative estimates are included). **c** Measures of various life-history and size-related traits in big-horn sheep (*Ovis canadensis*; based on Reale and Festa-Bianchet 2000; Coltman et al. 2005) and domestic sheep (*O. aries*; based on Milner et al. 2000; Coltman et al. 2001; Beraldi et al. 2007). Correlation for Big horn was −0.22 ± 0.16, and for domestic sheep it was −0.35 ± 0.18. **d** Measures of linear size measures from the genus *Drosophila* (based on many studies). The correlation is −0.10 ± 0.06

traits is about 1%, while for linear size traits it is about 0.1%, but the heritabilities of the linear size traits are twice those of the life-history traits. If size is measured as weight, however, we found much higher evolvabilities around 1%. This is not surprising, because mean-scaled variances depend on dimension. The general relationship of mean-scaled variances to dimension is complicated and depends on both the distribution and correlation of the dimensions, but two special-case low-variance approximations are illuminating. If the dimensions are perfectly correlated, then mean-scaled variances scale approximately with the square of the dimension making us expect weights and volumes to have nine times higher evolvabilities than linear measurements. The other extreme is when the dimensions are independent, then mean-scaled variances scale with dimension, making weights and volumes have three times the value of linear measurements. In our data the relation between weights and linear measurements is close to the former situation suggesting strongly correlated dimensions. Our sample size for area measures is too sparse to add to this discussion.

Table 1 Median evolvabilities and heritabilities for different trait categories. Size 1, 2, 3 and # refers to size measurement on linear, quadratic, cubic and count scales, respectively

Trait	Additive			Whole data set		
	$e_\mu = I_A$	h^2	n	$e_\mu = I_A$	h^2	n
Total	0.26 ± 0.03%	0.29 ± 0.01	960	0.36 ± 0.03%	0.30 ± 0.01	1,465
Size (1)	0.09 ± 0.01%	0.35 ± 0.02	394	0.12 ± 0.01%	0.35 ± 0.01	542
Size (2)	0.05 ± 0.21%	0.15 ± 0.06	12	0.72 ± 0.35%	0.20 ± 0.03	27
Size (3)	0.94 ± 0.18%	0.27 ± 0.02	125	1.23 ± 0.26%	0.28 ± 0.02	159
Size (#)	0.21 ± 0.11%	0.40 ± 0.07	50	0.39 ± 0.13%	0.44 ± 0.04	64
Life hist.	0.95 ± 0.15%	0.16 ± 0.04	149	1.04 ± 0.17%	0.20 ± 0.02	230
Physio.	0.49 ± 0.14%	0.12 ± 0.05	89	0.66 ± 0.24%	0.20 ± 0.04	129
Growth	1.81 ± 2.19%	0.32 ± 0.20	6	0.42 ± 0.30%	0.16 ± 0.02	30
Behavior	1.93 ± 0.46%	0.26 ± 0.07	44	1.56 ± 0.39%	0.23 ± 0.06	53
Shape	0.37 ± 0.17%	0.29 ± 0.05	42	0.42 ± 0.15%	0.27 ± 0.03	80
Pattern	1.68 ± 1.11%	0.47 ± 0.09	13	1.22 ± 0.32%	0.53 ± 0.03	103
Dev. st.	0.05 ± 0.21%	0.03 ± 0.02	28	0.38 ± 0.26%	0.03 ± 0.01	34
Houle92	0.25 ± 0.02%	0.33 ± 0.01	842	–	–	–

The cubic scale subset contains weight measurements with a small number of volume measurements. The "additive" columns contain estimates of additive variance that has some level of control for dominance and epistatic variance, as well as realized heritabilities, and thus gives the most appropriate measures of evolvability. The Houle92 row gives the corresponding statistics from the data of Houle (1992), who only included narrow-sense estimates of additive variance

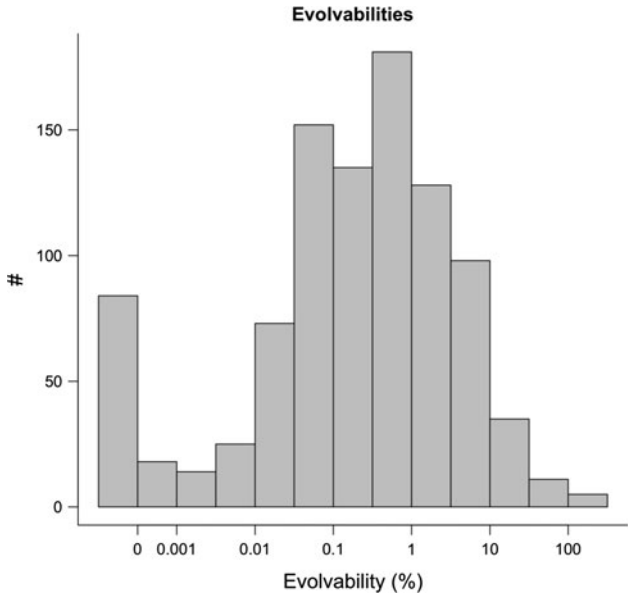

Fig. 5 Histogram of evolvabilities: Mean-scaled additive genetic variances plotted on a $\log_{10}$ scale. The leftmost bin include negative and zero estimates. The median is 0.26 ± 0.03%. This plot is restricted to "narrow sense" and realized evolvabilities, and does not include estimates based on full-sib or clonal variance

Supplementary figures 1–4 show that our results are not qualitatively dependent on including full-sib and broad-sense estimates of "heritability". The R^2 between mean-scaled and variance-scaled additive variances based on narrow-sense estimates alone was 0.7 ± 0.3%, slightly less than in the full data set. Within full-sib and broad-sense estimates alone the R^2 were slightly higher at 5 ± 2% and 2 ± 2%, respectively.

Finally, Fig. 5 shows that evolvabilities range over several orders of magnitude.

Why are Heritability and Evolvability Not Correlated?

Since both heritability and evolvability are measures of additive genetic variation, the absence of a correlation may seem surprising. We need to understand why this is so. Our working hypothesis is that there exist strong functional and statistical relations between the additive genetic variance and other variance components. These relations produce positive correlations (across populations and traits) between the variance components that tend to cancel the expected positive relation between additive variance and heritability.

Dominance and epistasis are traditionally conceptualized as independent of additive effects, but this is misleading. Consider simple additive-by-additive epistasis resulting from interactions of pairs of alleles at different loci. In a model with additive and additive-by-additive epistatic variance, the heritability is

$$h^2 = \frac{V_A}{V_A + V_{AA} + V_E}, \tag{8}$$

where V_{AA} is the additive-by-additive epistatic variance, and V_E is environmental or residual variance. Clearly, a correlation between V_A and V_{AA} could obscure an otherwise positive relation between V_A and h^2, and

indeed there are strong theoretical reasons to expect such a correlation. In fact, using the multilinear model of gene interactions, Hansen and Wagner (2001) showed that there exists a scaling relationship of the following form

$$V_{AA} = \varepsilon_e^2 V_A^2/2, \tag{9}$$

where ε_e^2 is an average of squared pairwise epistasis coefficients. A pairwise epistasis coefficient in the multilinear model determines how much a gene substitution on one locus will change the effect of a gene substitution at another locus. If the genotype-phenotype map (the phenotypic effects of changes in the genotype) stays constant, but the additive variance is changed (by changing allele frequencies) this equation predicts that the additive-by-additive epistatic variance will increase with the square of the additive variance. In other words, the heritability is

$$h^2 = \frac{V_A}{V_A + \varepsilon_e^2 V_A^2/2 + V_E}, \tag{10}$$

which turns into a decreasing function of additive variance when the additive variance becomes sufficiently large (Fig. 6). Hansen and Wagner (2001) further showed that a k-order epistatic variance component will scale with the k'th power of the additive variance. This should not be surprising if we think of additive and epistatic effects as first- and higher-order approximations of the genotype-phenotype map. When there is little variation, the first-order approximation is good, and almost all the genetic variance is additive, but as we increase allelic variation, second- and higher-order terms become relatively more important such that the fraction of variance described by the first-order effects eventually goes down even when the absolute amount of variance goes up. An added complication is that most estimates of V_A also include some contribution from V_{AA}, which will further complicate the relationship between h^2 and V_A.

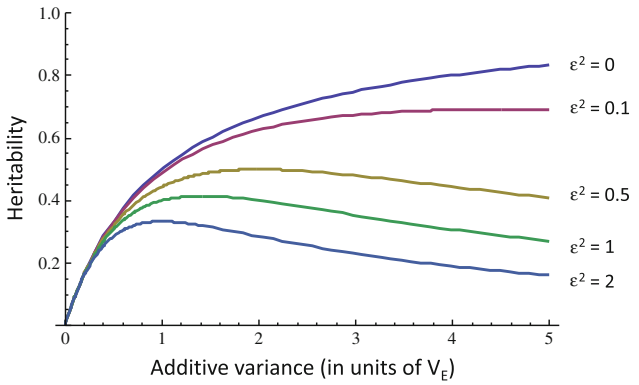

Fig. 6 Relationship between heritability and additive genetic variance in the multilinear epistatic model: On the x-axis is additive genetic variance in units of environmental variance. The parameter ε^2 is a measure of the strength of pairwise epistasis (Hansen and Wagner 2001; also called θ^2 in Carter et al. 2005). In the absence of epistasis ($\varepsilon^2 = 0$), heritability is a concave function of additive variance. In the presence of epistasis, the heritability will reach a maximum and then decline as additive variance is increased

There are similar, albeit more erratic, relations between additive and dominance variance. We can illustrate this with a single diallelic locus. Here,

$$V_A = 2pq(a + d(q - p))^2, \tag{11a}$$

$$V_D = (2pqd)^2, \tag{11b}$$

where the trait values of the two homozygotes and the heterozygote are −a, a, and d, respectively, and the two alleles have frequencies p and q. In Fig. 7a we show the complex relation between V_A and V_D for three levels of dominance. In the absence of other sources of variance, this makes the relation between heritability and additive variance strongly nonlinear and with little predictive value as illustrated in Fig. 7b.

For many researchers the intuition behind variance scaling may rest on the assumption that environmental variance is independent of the genetic variances. This assumption, embedded in theoretical population genetics, where environmental variance is routinely treated as an invariant parameter, and where predictions about additive variance and heritability are not distinguished, is, however, highly problematic. There are many biological reasons why genetic and environmental variation should be correlated. The most fundamental is perhaps that they are both expected to depend on the complexity of the character. A character with many interrelated parts or complex development has many potential targets where both genetic and environmental perturbation can act (Houle 1992, 1998, 2001), and a character with a sensitive development will be similarly vulnerable to both genetic and environmental perturbations. We also note that theoretical models tend to predict that genetic and environmental canalization may happen under similar circumstances, and the major driver of genetic canalization may in fact be its link with environmental canalization (Wagner et al. 1997; de Visser et al. 2003). Finally, stabilizing selection can not distinguish between components of variance and will reduce them all proportionally within a generation. Hence, regardless of genetic architecture, stabilizing selection will act directly to reduce additive variance, but leave heritability unaffected.

In Fig. 8 we plot mean-scaled additive variances against mean-scaled residual variances computed as $I_E = I_A(1 - h^2)/h^2$. This I_E contains both environmental variation and non-additive components of genetic variation, but we choose to use this rather than a direct estimate of environmental variation, as rather few studies achieve a good separation of residual genetic and environmental variance.

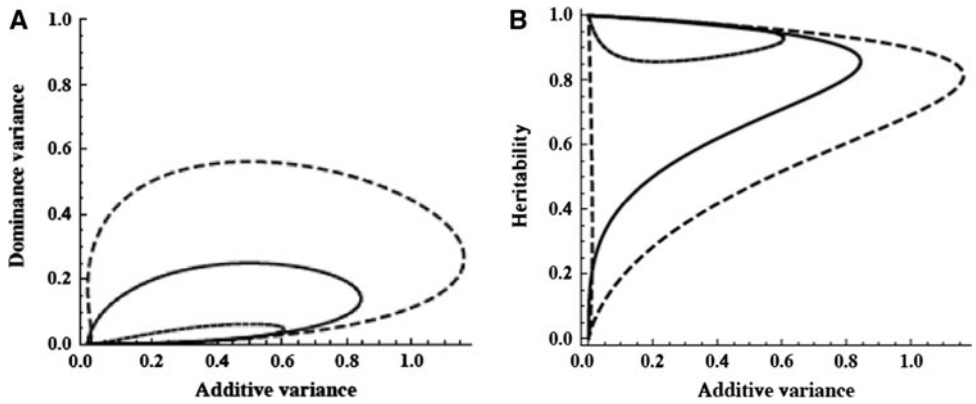

Fig. 7 Relationship between additive and dominance variance in a one-locus two-allele model: **a** The relationship between additive variance and dominance variance. **b** The relationship between additive variance and heritability. The three graphs in each figure are parametric plots for allele frequencies ranging from 0 to 1 based on equations in main text with a = 1. The *solid line* is for complete dominance (d = 1), the densely *dashed line* is for intermediate dominance (d = 0.5), and the less densely *dashed line* is for overdominance (d = 1.5)

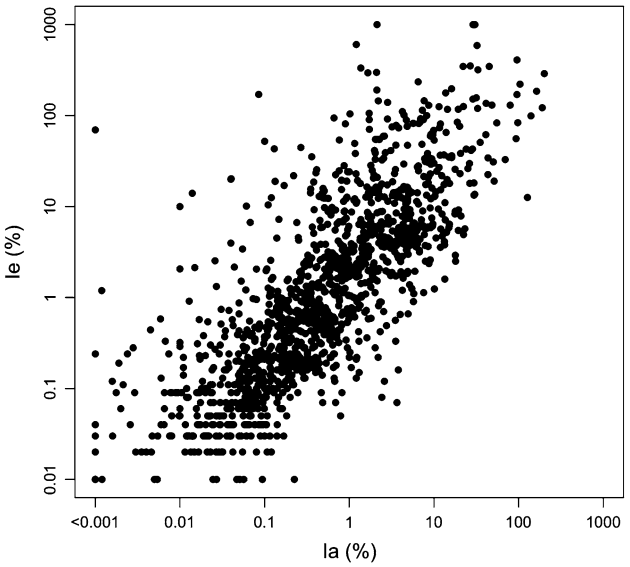

Fig. 8 Relationship between mean-scaled residual variance (computed as $I_e = I_A(1 - h^2)/h^2$) and mean-scaled additive variance for the entire data set. The correlation is 0.77 ± 0.01 (on the shown $\log_{10}$ scale). Pairs with nonpositive values for either variable excluded

In any case, the data in Fig. 8 support a positive relationship between additive variance and the residual variance with an $R^2 = 60 \pm 2\%$ for $\log_{10}$-transformed data. This relationship is similar within all trait categories (e.g. $R^2 = 59 \pm 4\%$ for linear size), with the single exception of developmental-stability traits where the R^2 is a mere $6 \pm 19\%$. Stirling et al. (2002) also found a correlation between $\log_{10}$ coefficients of additive variance and residual variance for behavioral traits. We note that correlations between non-log-transformed I_A and I_e are much weaker ($R^2 = 6 \pm 5\%$), because they are dominated by poorly correlated large values. Restricting the ranges to between 0 and 2% gives $R^2 = 28 \pm 4\%$. Heritabilities are negatively correlated with I_e (correlation = -0.17 ± 0.02 on I_e and -0.37 ± 0.02 on $\log_{10}[I_e]$).

Finally, we must consider the effect of scaling relations between the mean and variance. Indeed, eliminating such relations is the usual motivation for mean scaling or log transformation. This is not the reason for mean scaling advocated here; instead we base our approach on a model of evolvability as proportional change, and this is logically, although perhaps not practically, independent of the scaling relation between mean and variance. Still, it is important to know this relation, and an "allometric" scaling relation between mean (squared) and variance may be one mechanism weakening the correlation between mean-scaled and variance-scaled entities. In our data the correlation of I_A with the mean was -0.13 ± 0.04 with both variables on log scale (and -0.004 ± 0.005 on original scale). Similar weak correlations were found for specific trait categories. These results, and similar results on the relation between CV_A and the mean in Houle (1992), show that evolvabilities are not numerically dependent on the mean, and that mean dependency can not explain the lack of a relation between evolvabilities and heritabilities (correlations between heritabilities and means are also close to zero).

Discussion

Standardization is done to eliminate numerical differences that are due to uninteresting variation in scale, size, and dimension. Scaling additive genetic variance with phenotypic variance is a natural and effective way of doing this, since we do expect all the genetic and environmental variance components to be similarly affected by scale, size, and dimension. Using the phenotypic variance as a measuring stick for additive variance, however, entails the

additional assumption that the other variance components do not vary systematically in relation to the additive variance for other reasons. If they do, then scaling additive variance with phenotypic variance becomes akin to a rubber scale that gets stretched when measuring something large. We have put forward a number of theoretical and empirical reasons to expect correlations of additive variance with dominance, epistatic and environmental variance components.

The consequence of these correlations is that evolvability and heritability are almost completely uncorrelated. This is true both in general and within most specific trait categories. It is particularly striking that the correlation is almost exactly zero even within linear morphological measurements. These results, and those of Houle (1992), imply that heritabilities and evolvabilities should not be used interchangably because there is almost no predictive power from one to the other. The only exception is that large heritability can usually be taken to imply non-zero evolvability.

Seemingly in contradiction to our results, Stirling et al. (2002) reported that heritabilities and additive genetic coefficients of variation were statistically significantly related for a large sample of behavioral traits. Statistical significance does not imply high predictive value, however, and Stirling et al. (2002) did not report whether there was a high R^2 between the two measures.

The empirical finding of no correlation between heritability and evolvability does not by itself provide any basis for choice between the two measures. In our view, however, heritability is not suitable as a stand-alone measure of evolutionary potential in natural populations. Heritability is useful as a predictor of evolutionary response to artificial selection when the selection differential can be controlled by the investigator, but not as a general predictor of response to natural selection that is not specifically measured in the same population. One reason for this is that the heritability is expected to be negatively correlated with its corresponding measures of selection (either selection differential or selection intensity) due to variation in levels of environmental variation (Hereford et al. 2004). Environmental variance has little effect on the response to directional selection, but as the heritability scales negatively and the selection differential scales positively with environmental variance, we expect a negative correlation between the two. Since levels of environmental variance are extremely variable across traits and populations, this could generate a strong negative correlation even with the relatively weak negative correlation between heritability and residual variation we found in our data. Another reason why heritabilities are dubious predictors of evolutionary potential emerges from the theoretically and empirically based expectation of strong correlation between the additive variance and the other variance components. This would cause a poor correlation between heritability and additive variance. Hence, we recommend that heritabilities not be used as measures of evolutionary potential in the context of natural selection.

This is far from current practice in the field of evolutionary biology, in which most researchers act as if they are unaware that scaling involves assumptions. Dividing trait values by their standard deviation before analysis is so routine that this practice seldom receives more than a brief mention in the methods. In many cases authors merely state that traits were standardized without specifying how. When they are aware of alternatives such as mean standardization, the most common approach is simply to report both alternatives, and it is rare to see a discussion of what the differences might mean. Within evolutionary quantitative genetics, it was particularly unfortunate that the foundational paper by Lande and Arnold (1983) recommended scaling traits with their standard deviation without discussing what assumptions this entails. The combined result is that most variance-scaled selection gradients and variance components reported in the literature are incompletely interpreted (see Hereford et al. 2004; Stinchcombe et al. 2008; Wilson 2008; Houle et al. 2011 for further criticism).

The unreflective use of variance scaling in general, and heritabilities in particular, has lead to some unwarranted generalizations and doubtless to many dubious specific conclusions. The striking example is the idea that life-history traits in general, and fitness components in particular, have little additive genetic variance and low evolutionary potential. Although such traits do tend to have low heritabilities, Houle's (1992) and our data show that they also tend to have high evolvabilities, which on a variance-standardized scale gets masked by the very high levels of environmental (or possibly non-additive genetic) variation in such traits (see also Price and Schluter 1991). Merilä and Sheldon (1999) argued that the low heritabilities of fitness traits were due to high levels of dominance and epistatic variance and not to low levels of additive variance, and Meffert et al. (2002) made a similar argument for behavioral traits. In our data however, as well as those of Houle et al. (1996), there is no indication that broad-sense estimates are much higher than narrow-sense estimates for these trait categories. Regardless of the exact role of non-additive variance, our data support the hypothesis that life-history traits and complex behavioral traits really do tend to have a high potential for rapid evolution, and this is most likely due to them being functions of many components through which many genes can induce variation.

In theoretical population genetics the environmental variance is almost always modeled as a single invariant parameter. Predictions about additive genetic variance are often phrased as predictions about heritabilities. For example, the "Lande—Turelli" debate on the maintenance of additive genetic variance was usually phrased in terms

of explaining observed heritabilities (e.g. Lande 1976; Turelli 1984; see Barton and Turelli 1989; Bürger 2000; Johnson and Barton 2005 for review). The technical aspect of this work is impeccable, but the assumption of a constant environmental variance unrelated to the additive variance is empirically unjustified. The robust result of this work is that additive genetic variances should increase with mutation rate and mutational target size, and decrease with strength of stabilizing selection, but none of these predictions are justified for heritabilities.

More generally, there appears to be no theoretical or empirical reason for rates of evolution or species divergence to depend on the level of heritability as long as the latter is positive. This may possibly explain why we still lack a robust body of tests of the hypothesis that character evolution is influenced by genetic constraints. The few studies that have found a relationship have been based on comparing very similar characters and/or used mean-scaled evolvabilities (e.g., Blows and Higgie 2003; Hansen et al. 2003b; Marroig and Cheverud 2005; McGuigan et al. 2005; Hunt 2007; Hansen and Houle 2008; Chenoweth et al. 2010; Grabowski et al. 2011).

Even if it is clear that heritabilities are often inappropriate measures of evolutionary potential, mean scaling also entails assumptions, and mean-scaled additive genetic variances should not be uncritically accepted as measures of evolvability. One important limitation is that mean scaling should not be used for traits on nominal, ordinal or interval scale types. Interval scales where there is no natural zero point are particularly important to note, since they are not uncommon for quantitative characters and since heritabilities are meaningful on these scales. We note, however, that some traits that seem to be on interval scales are really on what Houle et al. (2011) termed signed ratio scales. These include traits such as residuals from a regression and signed asymmetries. In these cases mean scaling can be based on average absolute values or on the mean of the trait they are derived from (as indeed was done in some of the studies in our data base). There are, however, genuine interval scales, such as time of year measured in calendar date, where no form of mean scaling is possible. We see no obvious general way to compare the evolvabilities of these with other traits. Proportions and fractions provide another difficulty. This is discussed in detail by Stinchcombe (2005), who notes that mean-scaling has the undesirable property that proportions that are inverses of each other will usually be assigned different evolvabilities. Evolvability of proportions is perhaps best studied on an unstandardized scale, but in some situations one could also compute their evolvability on an assumed underlying ratio scale. A final issue with mean-standardized variances is that they are not invariant to dimension (Lande 1977; Houle 1992). As explained above, the effect of dimension depends on what statistical dependencies hold among the dimensional axes. If one is willing to assume that the dimensional axes are independent, for example, one could easily correct for dimension by dividing the evolvabilities with the dimension of the trait (in which case they would all be expressed as evolvabilities on a one-dimensional scale). We do suggest, however, that the higher evolvability of higher-dimensional traits could be seen as a real biological phenomenon. There simply are more ways to change a higher-dimensional trait.

More generally we can say that mean-scaled evolvabilities measure the potential for proportional change, and are meaningful whenever proportional changes are meaningful. Situations in which proportional changes may be less meaningful include tiny or rudimentary traits and traits with extremely skewed distributions. Apart from the statistical concern that means in such situations may have large relative errors, one can easily get large evolvabilities because it is easy to make large proportional changes, but this does not mean that the trait is highly evolvable in absolute size.

Fundamentally, standardization should only be used when the point is to compare qualitatively different traits. If the goal is to assess the evolvability of a single trait on its own terms, this should always be done on the original scale; i.e. on a scale that derives its meaning from the properties of the trait, the question, and the context.

Our purpose with this article has been to study the relationship between evolvability and heritability, and not to study evolvabilities per se. Still, we note that our results support the notion that quantitative traits are often highly evolvable when taken in isolation over short time spans. An evolvability of 0.1%, which is about the median for linear morphological traits in our database, may sound small. It predicts a tenth of a percent change per generation for traits under unit selection (Hansen et al. 2003b). Changes of this size may be hard to detect over a single generation, but can generate large changes over a few hundreds of generations. To a first approximation, the percent change over t generations with an evolvability of e_μ and strength of selection β_μ is $(1 + e_\mu\beta_\mu)^t$, and with $e_\mu = 0.001$ and $\beta_\mu = 1$ this will make a 10% change in slightly less than 100 generations, and a doubling of the trait in about 700 generations. With $\beta_\mu = 0.3$, the median bias-corrected estimate of multivariate gradients from the meta analysis of Hereford et al. (2004), those figures would be 317 and 2,311 generations, respectively. Thus, a large majority of our estimates indicate that the typical strong directional selection observed in nature is perfectly capable of generating large "qualitative" changes in the traits on less than a geological time scale. This seems to indicate that a lack of evolvability would rarely be a constraint on macroevolution.

This conclusion may, however, be premature (Hansen and Houle 2004; Blows and Hoffmann 2005; Blows 2007; Blows

and Walsh 2009; Kirkpatrick 2009; Walsh and Blows 2009). First, 7% of all our e_μ estimates were zero or negative, and this may be an underestimate as Palmer (2000) has documented a publication bias against negative and "nonsignificant" heritabilities. Second, about 9% of our 1,358 non-negative estimates were less than $e_\mu = 0.001\%$, which means that it would take more than 70,000 generations to double the trait under unit selection. We can therefore not rule out that direct genetic constraints are important for a substantial minority of traits. Measures of developmental stability would be one category in which direct genetic constraints could be common. We caution, however, that small evolvability estimates also have large relative errors, and it is therefore difficult to conclusively demonstrate genetic constraints. Third, univariate evolvabilities do not include the effects of genetic correlations and pleiotropic constraints. Such constraints can be formalized as conditional evolvabilities, the evolvability of a trait that is left when other traits remain fixed under stabilizing selection (Hansen 2003; Hansen et al. 2003a, b; Hansen and Houle 2008; Walsh and Blows 2009; see also Kirkpatrick 2009). Studies of conditional evolvabilities show that constraints generated by even a few other traits can reduce evolvabilities dramatically (Hansen et al. 2003a, b; Jensen et al. 2003; Parker and Garant 2004; Rolff et al. 2005; Rønning et al. 2007; Hansen and Houle 2008; Kirkpatrick 2009; McGuigan and Blows 2010; Grabowski et al. 2011). Considering that the genes affecting any focal trait may be subject to a large number of selective constraints, the hypothesis that many traits have very low conditional evolvability deserves serious attention (e.g. Walsh and Blows 2009). A final factor to consider is that evolvabilities are evolvable. There are two aspects to this. The first is that additive genetic variance will change during a response to selection due to allele-frequency changes, and that this eventually leads to an exhaustion of evolvability. Mutational evolvability, the mean-scaled mutational variance generated per generation, may therefore be a more relevant measure of long-term evolutionary potential (Houle et al. 1996; Hansen and Houle 2004). The second aspect is that the allelic effects themselves may change due to epistatic interactions with the evolving genetic background (Hansen and Wagner 2001). If there is a systematic tendency for allele substitutions in the direction of selection to reduce the effect of other such allele substitutions, then we would see a rapid reduction of both standing and mutational evolvability as a side effect of selection on the trait (Carter et al. 2005; Hansen et al. 2006). Hansen and Houle (2004) called this an epistatic constraint.

Heritabilities are not only used as measures of evolutionary potential. Our criticisms have been concerned with this use and do not necessarily imply that heritabilities are not useful for other purposes. We note, however, that the perils of using a rubber scale may also show in other contexts. The concept of heritability is commonly used in human genetics, and estimates from highly standardized sub-population samples are sometimes naively extrapolated to the population at large, or even to argue that differences between social, ethnic or temporal groups are genetically based. Such extrapolations have been criticized (e.g. Layzer 1974; Feldman and Lewontin 1975; Gould 1981; Jacoby and Glauberman 1995). Here, we just note that a failure (or unwillingness) to consider scale as a dynamic entity may be one reason why such extrapolations are not immediately recognized as problematic. The within-population heritability is relative to within-population variance, and if a high proportion of this is genetic it does not mean a high proportion of the variance in a larger population or between groups would also be genetic. Furthermore, observing a high heritability does not imply a high level of genetic determination, or that the trait is insensitive to environmental factors. It could simply mean that the environment in which the heritability was measured was relatively stable. A variance scale is population and environment specific and not a fixed entity. More generally we note that all the concepts we have discussed here, heritabilities, evolvabilities, variance components, and means are population variables that describe the properties of populations. They are not applicable to individuals, and not transferable from one population or environment to another.

Our main conclusion is that heritabilities should not be used as measures of evolutionary potential in natural populations. Beyond this, our findings illustrate how the choice of scale is a fundamental and extremely important methodological decision in evolutionary studies. Naive standardizations and the general absence of scale considerations in the interpretation and discussion of results is a serious problem in evolutionary biology and related fields (Houle et al. 2011). Scaling is simply a missing element of the methodological foundation in these fields, and until it becomes an integral part of model building and argumentation, this methodological omission will continue to generate serious systemic errors.

Acknowledgments We thank the many participants in the measurement-theory discussion group at the University of Oslo during 2008 and 2009 for discussions of the topic of this paper. DH was supported in this research by a visiting professorship at CEES and TFH by grant #177857 from the Norwegian Research Council.

Appendix

List of papers used in the literature review: Agrawal et al. (2002), Arnold and Phillips (1999), Ashman (2003), Asteles et al. (2006), Bacigalupe et al. (2004), Beraldi et al. (2007), Berwaerts et al. (2008), Birkhead et al. (2006),

Boake and Konigsberg (1998), Brandt and Greenfield (2004), Brodie (1993), Bryant and Meffert (1995), Cadee (2000), Campbell (1996, 1997), Caruso (2004), Caruso et al. (2005), Charmantier et al. (2004a, b), Cheetham et al. (1993, 1994), Cheverud (1996), Coltman et al. (2001, 2005), Conner and Via (1993), Conner et al. (2003, Cotter et al. 2004), Czesak and Fox (2003a, b), De Winter (1992), Evanno et al. (2006), Evans and Marshall (1996), Evans et al. (2006), Fenster and Carr (1997), Fernandez et al. (2003), Fox (1993), Fox et al. (1999), Friberg et al. (2005), Garant et al. (2004), Garcia-Gonzales and Simmons (2005), Gardner and Latta (2008), Gomez et al. (2009), Gomez-Mestre et al. (2008), Gray and Cade (1999), Groeters and Dingle (1996), Han and Lincoln (1994), Hansen et al. (2003b), Hawthorne (1997), Hegyi et al. (2002), Hendrickx et al. (2008), Hoffman et al. (2006), Hoffmann and Schiffer (1998), Horne and Ylönen (1998), House and Simmons (2005), House et al. (2008), Hughes (1995), Ivy (2007), Jensen et al. (2003, 2008), Jia et al. (2000), Johnson et al. (2009), Juenger and Bergelson (2000), Kaczorowski et al. (2008), Karoly and Conner (2000), Kause et al. (1999, 2001), Kellermann et al. (2006), Ketola and Kotiaho (2009), Kilpimaa et al. (2005), Kobayashi et al. (2003), Koelwijn and Hunscheid (2000), Kontiainen et al. (2008), Kruuk et al. (2002, 2003), Larsson (1993), Lauteri et al. (2004), Le Galliard et al. (2006), Leamy (1999), Lew et al. (2006), Linder and Rice (2005), Long et al. (2009), Lynch et al. (1999), MacColl and Hatchwell (2003), Magalhaes et al. (2007), Manier et al. (2007), Mappes and Koskela (2004), Mazer et al. (1999), McAdam and Boutin (2003), Merilä (1997), Merilä and Gustafsson (1993), Merilä et al. (1998), Messina and Fry (2003), Miller and Sinervo (2007), Milner et al. (2000), Morrow et al. (2008), Nespolo et al. (2003, 2005), Nilsson et al. (2009), Noach et al. (1996), ONeil (1997), Ostrowski et al. (2002), Parker and Garant (2004), Pélabon et al. (2004), Pelletier et al. (2007), Perez and Garcia (2002), Perry et al. (2004), Pettay et al. (2008), Platenkamp and Shaw (1992), Platenkamp and Shaw (1993), Podolsky et al. (1997), Polak and Starmer (2001), Radwan (1998), Rauter and Moore (2002), Reale and Festa-Bianchet (2000), Reale et al. (2003), Ritland and Ritland (1996), Rodriguez and Greenfield (2003), Roff (1995), Rolff et al. (2005), Rønning et al. (2007), Routley and Husband (2005), Ryder and Siva-Jothy (2001), Sakai et al. (2008), Santos (1996, 2001, 2002), Santos et al. (1992), Sarkissian and Harder (2001), Sgro and Hoffmann (1998), Shaw and Platenkamp (1993), Shaw et al. (1995), Sherrard et al. (2009), Simmons (2003), Simmons and Garcia-Gonzalez (2007), Simmons and Kotiaho (2002), Simons and Johnston (2006), Simons and Roff (1994), Simons et al. (1998), Steigenga et al. (2005), Teplitsky et al. (2009), Theriault et al. (2007), Thessing and Ekman (1994), Thiede (1998), Thomas and Simmons (2008), Tonsor and Goodnight (1997), Tucic and Stojkovic (2001), Van Kleunen and Ritland (2004), Verhoeven et al. (2004), Watkins (2001), Wayne et al. (1997), Weber and Scheiner (1992), Weller et al. (2006), Wilson et al. (2003), Wilson et al. (2005), Windig (1994), Winn (2004), Woods et al. (1998), Worley and Barrett (2000).

References

Agrawal, A. A., Conner, J. K., Johnson, M. T. J., & Wallsgrove, R. (2002). Ecological genetics of an induced plant defence against herbivores: Additive genetic variance and costs of phenotypic plasticity. *Evolution, 56*, 2206–2213.

Agrawal, A. F., & Stinchcombe, J. R. (2009). How much do genetic covariances alter the rate of adaptation. *Proceedings of the Royal Society B: Biological Sciences, 276*, 1182–1191.

Arnold, S. J., Pfrender, M. E., & Jones, A. G. (2001). The adaptive landscape as a conceptual bridge between micro- and macro-evolution. *Genetica, 112*(113), 9–32.

Arnold, S. J., & Phillips, P. C. (1999). Hierarchical comparison of genetic variance-covariance matrices. II. Coastal-inland divergence in the garter snake, *Thamnophis elegans*. *Evolution, 53*, 1516–1527.

Ashman, T. L. (2003). Constraints on the evolution of males and sexual dimorphism: Field estimates of genetic architecture of reproductive traits in three populations of gynodioecious *Fragaria virginiana*. *Evolution, 57*, 2012–2025.

Asteles, P. A., Moore, A. J., & Preziosi, R. F. (2006). A comparison of methods to estimate cross-environment genetic correlations. *Journal of Evolutionary Biology, 19*, 114–122.

Bacigalupe, L. D., Nespolo, R. F., Bustamante, D. M., & Bozinovic, F. (2004). The quantitative genetics of sustained energy budget in a wild mouse. *Evolution, 58*, 421–429.

Barton, N. H., & Turelli, M. (1989). Evolutionary quantitative genetics: how little do we know? *Annual Review of Genetics, 23*, 337–370.

Beraldi, D., Mcrae, A. F., Gratten, J., Slate, J., Visscher, P. M., & Pemberton, J. M. (2007). Mapping quantitative trait loci underlying fitness-related traits in a free-living sheep population. *Evolution, 61*, 1403–1416.

Berwaerts, K., Matthysen, E., & Van Dyck, H. (2008). Take-off flight performance in the butterfly *Pararge aegeria* relative to sex and morphology: A quantitative genetic assessment. *Evolution, 62*, 2525–2533.

Birkhead, T. R., Pellatt, E. J., Matthews, I. M., Roddis, N. J., Hunter, F. M., McPhie, F., et al. (2006). Genic capture and the genetic basis of sexually selected traits in the zebra finch. *Evolution, 60*, 2389–2398.

Blows, M. W. (2007). A tale of two matrices: Multivariate approaches in evolutionary biology. *Journal of Evolutionary Biology, 20*, 1–8.

Blows, M. W., Chenoweth, S. F., & Hine, E. (2004). Orientation of the genetic variance-covariance matrix and the fitness surface for multiple male sexually selected traits. *The American Naturalist, 163*, 329–340.

Blows, M. W., & Higgie, M. (2003). Genetic constraints on the evolution of mate recognition under natural selection. *The American Naturalist, 161*, 240–253.

Blows, M. W., & Hoffmann, A. A. (2005). A reassessment of genetic limits to evolutionary change. *Ecology, 86*, 1371–1384.

Blows, M. W., & Walsh, B. (2009). Spherical cows grazing in flatland: Constraints to selection and adaptation. In J.

Vanderwerf, H. U. Graser, R. Frankham, & C. Gondro (Eds.), *Adaptation and fitness in animal populations – evolutionary and breeding perspectives on genetic resource management*. Berlin: Springer.

Boake, C. R. B., & Konigsberg, L. (1998). Inheritance of male courtship behavior, aggressive success, and body size in *Drosophila silvestris*. *Evolution, 52*, 1487–1492.

Brandt, L. S. E., & Greenfield, M. D. (2004). Condition-dependent traits and the capture of genetic variance in male advertisement song. *Journal of Evolutionary Biology, 17*, 821–828.

Brodie, E. D. (1993). Homogeneity of the genetic variance-covariance matrix for antipredator traits in 2 natural-populations of the garter snake *Thamnophis ordinoides*. *Evolution, 47*, 844–854.

Brookfield, J. F. Y. (2009). Evolution and evolvability: Celebration Darwin 200. *Biology Letters, 5*, 44–46.

Bryant, E. H., & Meffert, L. M. (1995). An analysis of selectional response in relation to a population bottleneck. *Evolution, 49*, 626–634.

Bürger, R. (2000). *The mathematical theory of selection, recombination, and mutation*. Chichester: Wiley.

Burton, G. W. (1952). Quantitative inheritance in grasses. *Proceedings of the Sixth International Grassland Congress, 6*, 277–283.

Cadee, N. (2000). Genetic and environmental effects on morphology and fluctuating asymmetry in nestling barn swallows. *Journal of Evolutionary Biology, 13*, 359–370.

Campbell, D. R. (1996). Evolution of floral traits in a hermaphroditic plant: Field measurements of heritabilities and genetic correlations. *Evolution, 50*, 1442–1453.

Campbell, D. R. (1997). Genetic and environmental variation in life-history traits of a monocarpic perennial: A decade-long field experiment. *Evolution, 51*, 373–382.

Carter, A. J. R., Hermisson, J., & Hansen, T. F. (2005). The role of epistatic gene interactions in the response to selection and the evolution of evolvability. *Theoretical Population Biology, 68*, 179–196.

Caruso, C. M. (2004). The quantitative genetics of floral trait variation in *Lobelia*: Potential constraints on adaptive evolution. *Evolution, 58*, 732–740.

Caruso, C. M., Maherali, H., Mikulyuk, A., Carlson, K., & Jackson, R. B. (2005). Genetic variance and covariance for physiological traits in Lobelia: Are there constraints on adaptive evolution? *Evolution, 59*, 826–837.

Charmantier, A., Kruuk, L. E. B., Blondel, J., & Lambrechts, M. M. (2004a). Testing for microevolution in body size in three blue tit populations. *Journal of Evolutionary Biology, 17*, 732–743.

Charmantier, A., Kruuk, L. E. B., & Lambrechts, M. M. (2004b). Parasitism reduces the potential for evolution in a wild bird population. *Evolution, 58*, 203–206.

Cheetham, A. H., Jackson, J. B. C., & Hayek, L. A. C. (1993). Quantitative genetics of bryozoan phenotypic evolution. 1. Rate tests for random change versus selection in differentiation of living species. *Evolution, 47*, 1526–1538.

Cheetham, A. H., Jackson, J. B. C., & Hayek, L. A. C. (1994). Quantitative genetics of bryozoan phenotypic evolution. 2. Analysis of selection and random change in fossil species using reconstructed genetic-parameters. *Evolution, 48*, 360–375.

Chenoweth, S. F., Rundle, H. D., & Blows, M. W. (2010). The contribution of selection and genetic constraints to phenotypic divergence. *The American Naturalist, 175*, 186–196.

Cheverud, J. (1996). Quantitative genetic analysis of cranial morphology in the cotton-top (*Saguinus oedipus*) and saddle-back (*S. fuscicollis*) tamarins. *Journal of Evolutionary Biology, 9*, 5–42.

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

Coltman, D. W., Pilkington, J., Kruuk, L. E. B., Wilson, K., & Pemberton, J. M. (2001). Positive genetic correlation between parasite resistance and body size in a free-living ungulate population. *Evolution, 55*, 2116–2125.

Conner, J. K., Franks, R., & Stewart, C. (2003). Expression of additive genetic variances and covariances for wild radish floral traits: Comparison between field and greenhouse environments. *Evolution, 57*, 487–495.

Conner, J., & Via, S. (1993). Patterns of phenotypic and genetic correlations among morphological and life-history traits in wild radish, *Raphanus raphanistrum*. *Evolution, 47*, 704–711.

Cotter, S. C., Kruuk, L. E. B., & Wilson, K. (2004). Costs of resistance: genetic correlations and potential trade-offs in an insect immune system. *Journal of Evolutionary Biology, 17*, 421–429.

Czesak, M. E., & Fox, C. W. (2003a). Evolutionary ecology of egg size and number in a seed beetle: Genetic trade-off differs between environments. *Evolution, 57*, 1121–1132.

Czesak, M. E., & Fox, C. W. (2003b). Genetic variation in male effects on female reproduction and the genetic covariance between the sexes. *Evolution, 57*, 1359–1366.

de Visser, J. A. G. M., Hermisson, J., Wagner, G. P., Ancel Meyers, L., Bagheri-Chaichian, H., Blanchard, J., et al. (2003). Evolution and detection of genetic robustness. *Evolution, 57*, 1959–1972.

De Winter, A. J. (1992). The gentic basis and evolution of acoustic mate recognition in a *Ribautodelphax* planthopper (homoptera, Delphacidae) 1. The female call. *Journal of Evolutionary Biology, 5*, 249–265.

Estes, S., & Arnold, S. J. (2007). Resolving the paradox of stasis: Models with stabilizing selection explain evolutionary divergence on all timescales. *The American Naturalist, 169*, 227–244.

Evanno, G., Castella, E., & Coudet, J. (2006). Evolutionary aspects of population structure for molecular and quantitative traits in the freshwater snail *Radix balthica*. *Journal of Evolutionary Biology, 19*, 1071–1072.

Evans, A. S., & Marshall, M. (1996). Developmental instability in *Brassica compestris* (Cruciferae): Fluctuating asymmetry of foliar and floral traits. *Journal of Evolutionary Biology, 9*, 717–736.

Evans, M. R., Roberts, M. L., Buchanan, K. L., & Goldsmith, A. R. (2006). Heritabilty of corticosterone response and changes in life history traits during selection in the zebra finch. *Journal of Evolutionary Biology, 19*, 343–352.

Falconer, D. S., & Mackay, T. F. C. (1996). *Introduction to quantitative genetics* (4th ed.). UK: Longman.

Feldman, M. W., & Lewontin, R. C. (1975). The heritability hang-up. *Science, 190*, 1163–1168.

Fenster, C. B., & Carr, D. E. (1997). Genetics of sex allocation in mimulus (Scrophulariaceae). *Journal of Evolutionary Biology, 10*, 641–661.

Fernandez, J., Rodriguez-Ramilo, S. T., Perez-Figueroa, A., Lopez-Fanjul, C., & Caballero, A. (2003). Lack of nonadditive genetic effects on early fecundity in *Drosophila melanogaster*. *Evolution, 57*, 558–565.

Fisher, R. A. (1951). Limits to intensive production in animals. *British Agriculture Bulletin, 4*, 217–218.

Fox, C. W. (1993). A quantitative genetic-analysis of oviposition preference and larval performance on 2 hosts in the bruchid beetle, *Callosobruchus maculatus*. *Evolution, 47*, 166–175.

Fox, C. W., Czesak, M. E., Mousseau, T. A., & Roff, D. A. (1999). The evolutionary genetics of an adaptive maternal effect: Egg size plasticity in a seed beetle. *Evolution, 53*, 552–560.

Friberg, U., Lew, T. A., Byrne, P. G., & Rice, W. R. (2005). Assessing the potential for an ongoing arms race within and between the sexes: Selection and heritable variation. *Evolution, 59*, 1540–1551.

Futuyma, D. J. (2010). Evolutionary constraint and ecological consequences. *Evolution, 64*, 1865–1884.

Garant, D., Sheldon, B. C., & Gustafsson, L. (2004). Climatic and temporal effects on the expression of secondary sexual characters: Genetic and environmental components. *Evolution, 58*, 634–644.

Garcia-Gonzales, F., & Simmons, L. W. (2005). The evolution of polyandry: Intrinsic sire effects contribute to embryo viability. *Journal of Evolutionary Biology, 18*, 1097–1103.

Gardner, K. M., & Latta, R. G. (2008). Heritable variation and genetic correlation of quantitative traits within and between ecotypes of *Avena barbata*. *Journal of Evolutionary Biology, 21*, 737–748.

Gomez, J. M., Abdelaziz, M., Munoz-Pajares, J., & Perfectti, F. (2009). Heritability and genetic correlation of corolla shape and size in *Erysimum mediohispanicum*. *Evolution, 63*, 1820–1831.

Gomez-Mestre, I., Touchon, J. C., Caccoccio, V. L., & Warkentin, K. M. (2008). Genetic variation in pathogen-induced early hatching of toad embryos. *Journal of Evolutionary Biology, 21*, 791–800.

Gomulkiewicz, R., & Houle, D. (2009). Demographic and genetic constraints on evolution. *The American Naturalist, 174*, E218–E229.

Gould, S. J. (1981). *The mismeasure of man*. Baskerville: Pelican Books.

Grabowski, M. W., Polk, J. D., & Roseman, C. C. (2011). Divergent patterns of integration and reduced constraint in the human hip and the origins of bipedalism. *Evolution, 65*, 1336–1356.

Gray, D. A., & Cade, W. H. (1999). Quantitative genetics of sexual selection in the field cricket, *Gryllus integer*. *Evolution, 53*, 848–854.

Groeters, F. R., & Dingle, H. (1996). Heritability of wing length in nature for the milkweed bug, *Oncopeltus fasciatus*. *Evolution, 50*, 442–447.

Hallgrimsson, B., Jamniczky, H., Young, N. M., Rolian, C., Parsons, T. E., Boughner, J. C., et al. (2009). Deciphering the palimpsest: Studying the relationship between morphological integration and phenotypic covariation. *Evolutionary Biology, 36*, 355–376.

Han, K. P., & Lincoln, D. E. (1994). The evolution of carbon allocation to plant secondary metabolites—a genetic-analysis of cost in *Diplacus aurantiacus*. *Evolution, 48*, 1550–1563.

Hansen, T. F. (2003). Is modularity necessary for evolvability? Remarks on the relationship between pleiotropy and evolvability. *Biosystems, 69*, 83–94.

Hansen, T. F. (2006). The evolution of genetic architecture. *Annual Review of Ecology, Evolution and Systematics, 37*, 123–157.

Hansen, T. F., Alvarez-Castro, J. M., Carter, A. J. R., Hermisson, J., & Wagner, G. P. (2006). Evolution of genetic architecture under directional selection. *Evolution, 60*, 1523–1536.

Hansen, T. F., Armbruster, W. S., Carlson, M. L., & Pélabon, C. (2003a). Evolvability and genetic constraint in *Dalechampia* blossoms: Genetic correlations and conditional evolvability. *Journal of Experimental Zoology, 296B*, 23–39.

Hansen, T. F., & Houle, D. (2004). Evolvability, stabilizing selection, and the problem of stasis. In M. Pigliucci & K. Preston (Eds.), *Phenotypic integration: Studying the ecology and evolution of complex phenotypes*. Oxford: Oxford University press.

Hansen, T. F., & Houle, D. (2008). Measuring and comparing evolvability and constraint in multivariate characters. *Journal of Evolutionary Biology, 21*, 1201–1219.

Hansen, T. F., Pélabon, C., Armbruster, W. S., & Carlson, M. L. (2003b). Evolvability and genetic constraint in *Dalechampia* blossoms: Components of variance and measures of evolvability. *Journal of Evolutionary Biology, 16*, 754–765.

Hansen, T. F., & Wagner, G. P. (2001). Modeling genetic architecture: A multilinear model of gene interaction. *Theoretical Population Biology, 59*, 61–86.

Hawthorne, D. J. (1997). Ecological history and evolution in a novel environment: Habitat heterogeneity and insect adaptation to a new host. *Evolution, 51*, 153–162.

Hegyi, G., Torok, J., & Toth, L. (2002). Qualitative population divergence in proximate determination of a sexually selected trait in the collared flycatcher. *Journal of Evolutionary Biology, 15*, 710–719.

Hendrickx, F., Maelfait, J. P., & Lens, L. (2008). Effect of metal stress on life history divergence and quantitative genetic architecture in a wolf spider. *Journal of Evolutionary Biology, 21*, 183–193.

Hendrikse, J. L., Parsons, T. E., & Hallgrimsson, B. (2007). Evolvability as the proper focus of evolutionary developmental biology. *Evolution & Development, 9*, 393–401.

Hereford, J., Hansen, T. F., & Houle, D. (2004). Comparing strengths of directional selection: How strong is strong? *Evolution, 58*, 2133–2143.

Hoffman, E. A., Mobley, K. B., & Jones, A. G. (2006). Male pregnancy and the evolution of body segmentation in seahorses and pipefishes. *Evolution, 60*, 404–410.

Hoffmann, A. A., & Schiffer, M. (1998). Changes in the heritability of five morphological traits under combined environmental stresses in *Drosophila melanogaster*. *Evolution, 52*, 1207–1212.

Horne, T. J., & Ylönen, H. (1998). Heritabilities of dominance-related traits in male bank voles (*Clethrionomys glareolus*). *Evolution, 52*, 894–899.

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

Houle, D. (1998). How should we explain variation in the genetic variance of traits? *Genetica, 102*(103), 241–253.

Houle, D. (2001). Characters as the units of evolutionary change. In G. P. Wagner (Ed.), *The character concept in evolutionary biology*. Massachusetts: Academic press.

Houle, D., Morikawa, B., & Lynch, M. (1996). Comparing mutational variabilities. *Genetics, 143*, 1467–1483.

Houle, D., Pélabon, C., Wagner, G. P., & Hansen, T. F. (2011). Measurement and meaning in biology. *The Quarterly Review of Biology, 86*, 3–34.

House, C. M., Evans, G. M. V., Smiseth, P. T., Stamper, C. E., Walling, C. A., & Moore, A. J. (2008). The evolution of repeated mating in the burying beetle, *Nicrophorus vespilloides*. *Evolution, 62*, 2004–2014.

House, C. M., & Simmons, L. W. (2005). The evolution of male genitalia: patterns of genetic variation and covariation in the genital sclerites of the dung beetle *Onthophagus taurus*. *Journal of Evolutionary Biology, 18*, 1281–1292.

Hughes, K. A. (1995). The evolutionary genetics of male life-history characters in *Drosophila melanogaster*. *Evolution, 49*, 521–537.

Hunt, G. (2007). Evolutionary divergence in directions of high phenotypic variance in the Ostracode genus *Poseidonamicus*. *Evolution, 61*, 1560–1576.

Ivy, T. M. (2007). Good genes, genetic compatibility and the evolution of polyandry: Use of the diallel cross to address competing hypotheses. *Journal of Evolutionary Biology, 20*, 479–487.

Jacoby, R., & Glauberman, N. (1995). *The bell curve debate: History, documents, opinions*. New York: Times Books.

Jensen, H., Sæther, B. E., Ringsby, T. H., Tufto, J., Griffith, S. G., & Ellegren, H. (2003). Sexual variation in heritability and genetic correlations of morphological traits in house sparrow (*Passer domesticus*). *Journal of Evolutionary Biology, 16*, 1296–1307.

Jensen, H., Steinsland, I., Ringsby, T. H., & Sæther, B. E. (2008). Evolutionary dynamics of a sexual ornament in the house sparrow (*Passer domesticus*): The role of indirect selection within and between sexes. *Evolution, 62*, 1275–1293.

Jia, F. Y., Greenfield, M. D., & Collins, R. D. (2000). Genetic variance of sexually selected traits in waxmoths: Maintenance by genotype × environment interaction. *Evolution, 54*, 953–967.

Johnson, M. T. J., Agrawal, A. A., Maron, J. L., & Salminen, J. P. (2009). Heritability, covariation and natural selection on 24 traits of common evening primrose (*Oenothera biennis*) from a field experiment. *Journal of Evolutionary Biology, 22*, 1295–1307.

Johnson, T., & Barton, N. (2005). Theoretical models of selection and mutation on quantitative traits. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 360*, 1411–1425.

Juenger, T., & Bergelson, J. (2000). The evolution of compensation to herbivory in scarlet gilia, *Ipomopsis aggregata*: Herbivore-imposed natural selection and the quantitative genetics of tolerance. *Evolution, 54*, 764–777.

Kaczorowski, R. L., Juenger, T. E., & Holtsford, T. R. (2008). Heritability and correlation structure of nectar and floral morphology traits in *Nicotiana alata*. *Evolution, 62*, 1738–1750.

Karoly, K., & Conner, J. K. (2000). Heritable variation in a family-diagnostic trait. *Evolution, 54*, 1433–1438.

Kause, A., Saloniemi, I., Haukioja, E., & Hanhimaki, S. (1999). How to become large quickly: Quantitative genetics of growth and foraging in a flush feeding lepidopteran larva. *Journal of Evolutionary Biology, 12*, 471–482.

Kause, A., Saloniemi, I., Morin, J. P., Haukioja, E., Hanhimaki, S., & Ruohomaki, K. (2001). Seasonally varying diet quality and the quantitative genetics of development time and body size in birch feeding insects. *Evolution, 55*, 1992–2001.

Kellermann, V. M., van Heerwaarden, B., Hoffmann, A. A., & Sgro, C. M. (2006). Very low additive genetic variance and evolutionary potential in multiple populations of two rainforest Drosophila species. *Evolution, 60*, 1104–1108.

Ketola, T., & Kotiaho, J. S. (2009). Inbreeding, energy use and condition. *Journal of Evolutionary Biology, 22*, 770–781.

Kilpimaa, J., Van de Casteele, T., Jokinen, I., Mappes, J., & Alatalo, R. V. (2005). Genetic and environmental variation in antibody and T-cell mediated responses in the great tit. *Evolution, 59*, 2483–2489.

Kirkpatrick, M. (2009). Patterns of quantitative genetic variation in multiple dimensions. *Genetica, 136*, 271–284.

Kobayashi, A., Tanaka, Y., & Shimada, M. (2003). Genetic variation of sex allocation in the parasitoid wasp *Heterospilus prosopidis*. *Evolution, 57*, 2659–2664.

Koelwijn, H. P., & Hunscheid, M. P. H. (2000). Intraspecific variation in sex allocation in hermaphroditic *Plantago coronopus* (L.). *Journal of Evolutionary Biology, 13*, 302–315.

Kontiainen, P., Brommer, J. E., Karell, P., & Petiainen, H. (2008). Heritability, plasticity and canalization of Ural owl egg size in a cyclic environment. *Journal of Evolutionary Biology, 21*, 88–96.

Kruuk, L. E. B., Slate, J., Pemberton, J. M., Brotherstone, S., Guinness, F., & Clutton-Brock, T. (2002). Antler size in red deer: Heritability and selection but no evolution. *Evolution, 56*, 1683–1695.

Kruuk, L. E. B., Slate, J., Pemberton, J. M., & Clutton-Brock, T. H. (2003). Fluctuating asymmetry in a secondary sexual trait: no associations with individual fitness, environmental stress or inbreeding, and no heritability. *Journal of Evolutionary Biology, 16*, 101–113.

Lande, R. (1976). Natural selection and random genetic drift in phenotypic evolution. *Evolution, 30*, 314–334.

Lande, R. (1977). On comparing coefficients of variation. *Systematic Zoology, 26*, 214–217.

Lande, R. (1979). Quantitative genetic analysis of multivariate evolution, applied to brain:body size allometry. *Evolution, 33*, 402–416.

Lande, R., & Arnold, S. J. (1983). The measurement of selection on correlated characters. *Evolution, 37*, 1210–1226.

Larsson, K. (1993). Inheritance of body size in the Barnacle Goose under different environmental conditions. *Journal of Evolutionary Biology, 6*, 195–208.

Lauteri, M., Pliura, A., Monteverdi, M. C., Brugnoli, E., Villani, F., & Eriksson, G. (2004). Genetic variation in carbon isotope discrimination in six European populations of *Castanea sativa* Mill. originating from contrasting localities. *Journal of Evolutionary Biology, 17*, 1286–1296.

Layzer, D. (1974). Heritability analysis of IQ scores: Science or numerology? *Science, 183*, 1259–1266.

Le Galliard, J. F., Massot, M., Landys, M. M., Meylan, S., & Clobert, J. (2006). Ontogenetic sources of variation is sexual size dimorphism in a viviparous lizard. *Journal of Evolutionary Biology, 19*, 690–704.

Leamy, L. (1999). Heritability of directional and fluctuating asymmetry for mandibular characters in random-bred mice. *Journal of Evolutionary Biology, 12*, 146–155.

Lew, T. A., Morrow, E. H., & Rice, W. R. (2006). Standing genetic variance for female resistance to harm from males and its relationship to intralocus sexual conflict. *Evolution, 60*, 97–105.

Linder, J. E., & Rice, W. R. (2005). Natural selection and genetic variation for female resistance to harm from males. *Journal of Evolutionary Biology, 18*, 568–575.

Long, T. A. F., Miller, P. M., Stewart, A. D., & Rice, W. R. (2009). Estimating heritability of female lifetime fecundity in a locally a adapted *Drosophila melanogaster* population. *Journal of Evolutionary Biology, 22*, 637–643.

Lynch, M., Pfrender, M., Spitze, K., Lehman, N., Hicks, J., Allen, D., et al. (1999). The quantitative and molecular genetic architecture of a subdivided species. *Evolution, 53*, 100–110.

Lynch, M., & Walsh, B. (1998). *Genetics and analysis of quantitative characters*. Massachusetts: Sinauer.

MacColl, A. D. C., & Hatchwell, B. J. (2003). Heritability of parental effort in a passerine bird. *Evolution, 57*, 2191–2195.

Magalhaes, S., Fayard, J., Janssen, A., Carbonell, D., & Olivieiri, I. (2007). Adaptation in a spider mite population after long-term evolution on a single host plant. *Journal of Evolutionary Biology, 20*, 2016–2027.

Manier, M. K., Seyler, C. M., & Arnold, S. A. (2007). Adaptive divergence within and between ecotypes of the terrestrial garter snake, *Thamnophis elegans*, assessed with Fst—Qst comparisons. *Journal of Evolutionary Biology, 20*, 1705–1719.

Mappes, T., & Koskela, E. (2004). Genetic basis of the trade-off between offspring number and quality in the bank vole. *Evolution, 58*, 645–650.

Marroig, G., & Cheverud, J. M. (2005). Size as a line of least evolutionary resistance: Diet and adaptive morphological radiation in New World monkeys. *Evolution, 59*, 1128–1142.

Marroig, G., Shirai, L. T., Porto, A., de Oliveria, F. B., & De Conto, V. (2009). The evolution of modularity in the mammalian skull II: Evolutionary consequences. *Evolutionary Biology, 36*, 136–148.

Mazer, S. J., Delesalle, V. A., & Neal, P. R. (1999). Responses of floral traits to selection on primary sexual investment in *Spergularia marina*: The battle between the sexes. *Evolution, 53*, 717–731.

McAdam, A. G., & Boutin, S. (2003). Effects of food abundance on genetic and maternal variation in the growth rate of juvenile red squirrels. *Journal of Evolutionary Biology, 16*, 1249–1256.

McGuigan, K., & Blows, M. W. (2010). Evolvability of individual traits in a multivariate context: Patitioning the additive genetic variance into common and specific components. *Evolution, 64*, 1899–1911.

McGuigan, K., Chenoweth, S. F., & Blows, M. W. (2005). Phenotypic divergence along lines of genetic variance. *The American Naturalist, 165*, 32–43.

Meffert, L. M., Hicks, S. K., & Regan, J. L. (2002). Nonadditive genetic effects in animal behavior. *The American Naturalist, 160*, S198–S213.

Merilä, J. (1997). Expression of genetic variation in body size of the collared flycatcher under different environmental conditions. *Evolution, 51*, 526–536.

Merilä, J., & Gustafsson, L. (1993). Inheritance of size and shape in a natural population of collared flycatchers, *Ficedula albicollis*. *Journal of Evolutionary Biology, 6*, 375–395.

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

Merilä, J., Sheldon, B. C., & Ellegren, H. (1998). Quantitative genetics of sexual size dimorphism in the collared flycatcher, *Ficedula albicollis*. *Evolution, 52*, 870–876.

Messina, F. J., & Fry, J. D. (2003). Environment-dependent reversal of a life history trade-off in the seed beetle *Callosobruchus maculatus*. *Journal of Evolutionary Biology, 16*, 501–509.

Miller, B. L. W., & Sinervo, B. (2007). Heritable body size mediates apparent life-history trade-offs in a simultaneous hermaphrodite. *Journal of Evolutionary Biology, 20*, 1554–1562.

Milner, J. M., Pemberton, J. M., Brotherstone, S., & Albon, S. D. (2000). Estimating variance components and heritabilities in the wild: A case study using the "animal model" approach. *Journal of Evolutionary Biology, 13*, 804–813.

Mitteroecker, P., & Bookstein, F. (2007). The conceptual and statistical relationship between modularity and morphological integration. *Systematic Biology, 56*, 818–836.

Mitteroecker, P., & Bookstein, F. (2009). The ontogenetic trajectory of the phenotypic covariance matrix, with examples from craniofacial shape in rats and humans. *Evolution, 63*, 727–737.

Morrow, E. H., Leijon, A., & Meerupati, A. (2008). Hemiclonal analysis reveals significant genetic, environmental and genotype x environment effects on sperm size in *Drosophila melanogaster*. *Journal of Evolutionary Biology, 21*, 1692–1702.

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

Nespolo, R. F., Bacigalupe, L. D., & Bozinovic, F. (2003). Heritability of energetics in a wild mammal, the leaf-eared mouse (*Phyllotis darwini*). *Evolution, 57*, 1679–1688.

Nespolo, R. F., Bustamante, D. M., Bacigalupe, L. D., & Bozinovic, F. (2005). Quantitative genetics of bioenergetics and growth-related traits in the wild mammal, *Phyllotis darwini*. *Evolution, 59*, 1829–1837.

Nilsson, J. Å., Åkesson, M., & Nilsson, J. F. (2009). Heritability of resting metabolic rate in a wild population of blue tits. *Journal of Evolutionary Biology, 22*, 1867–1874.

Noach, E. J. K., de Jong, G., & Scharloo, W. (1996). Phenotypic plasticity in morphological traits in two populations of *Drosophila melanogaster*. *Journal of Evolutionary Biology, 9*, 831–844.

ONeil, P. (1997). Natural selection on genetically correlated phenological characters in *Lythrum salicaria* L (Lythraceae). *Evolution, 51*, 267–274.

Ostrowski, M. F., Jarne, P., & David, P. (2002). A phallus for free? Quantitative genetics of sexual trade-offs in the snail *Bulinus truncatus*. *Journal of Evolutionary Biology, 16*, 7–16.

Palmer, A. R. (2000). Quasireplication and the contract of error: Lessons from sex ratios, heritabilities and fluctuating asymmetry. *Annual Review of Ecology and Systematics, 31*, 441–480.

Parker, T. H., & Garant, D. (2004). Quantitative genetics of sexually dimorphic traits and capture of genetic variance by a sexually-selected condition-dependent ornament in red junglefowl (*Gallus gallus*). *Journal of Evolutionary Biology, 17*, 1277–1285.

Pavlicev, M., Cheverud, J. M., & Wagner, G. P. (2009). Measuring morphological integration using eigenvalue variance. *Evolutionary Biology, 36*, 157–170.

Pélabon, C., Hansen, T. F., Carlson, M. L., & Armbruster, W. S. (2004). Variational and genetic properties of developmental stability in *Dalechampia scandens*. *Evolution, 58*, 504–514.

Pelletier, F., Reale, D., Garant, D., Coltman, D. W., & Festa-Bianchet, M. (2007). Selection on heritable seasonal phenotypic plasticity of body mass. *Evolution, 61*, 1969–1979.

Perez, A., & Garcia, C. (2002). Evolutionary responses of *Drosophila melanogaster* to selection at different larval densities: Changes in genetic variation, specialization and phenotypic plasticity. *Journal of Evolutionary Biology, 15*, 524–536.

Perry, G. M. L., Audet, C., Laplatte, B., & Bernatchez, L. (2004). Shifting patterns in genetic control at the embryo-alevin boundary in brook charr. *Evolution, 58*, 2002–2012.

Pettay, J. E., Charmantier, A., Wilson, A. J., & Lummaa, V. (2008). Age-specific genetic and maternal effects in fecundity of preindustrial Finnish women. *Evolution, 62*, 2297–2304.

Platenkamp, G. A. J., & Shaw, R. G. (1992). Environmental and genetic constraints on adaptive population differentiation in *Anthoxanthum odoratum*. *Evolution, 46*, 341–352.

Platenkamp, G. A. J., & Shaw, R. G. (1993). Environmental and genetic maternal effects on seed characters in *Nemophila menziesii*. *Evolution, 47*, 540–555.

Podolsky, R. H., Shaw, R. G., & Shaw, F. H. (1997). Population structure of morphological traits in *Clarkia dudleyana.* II. Constancy of within-population genetic variance. *Evolution, 51*, 1785–1796.

Polak, M., & Starmer, W. T. (2001). The quantitative genetics of fluctuating asymmetry. *Evolution, 55*, 498–511.

Polly, D. (2008). Developmental dynamics and G-matrices: Can morphometric spaces be used to model phenotypic evolution? *Evolutionary Biology, 35*, 83–96.

Price, T., & Schluter, D. (1991). On the heritability of life-history traits. *Evolution, 45*, 853–861.

Radwan, J. (1998). Heritability of sperm competition success in the bulb mite, *Rhizoglyphus robini*. *Journal of Evolutionary Biology, 11*, 321–327.

Rauter, C. M., & Moore, A. J. (2002). Evolutionary importance of parental care performance, food resources, and indirect genetic effects in a burying beetle. *Journal of Evolutionary Biology, 15*, 407–417.

Reale, D., Berteaux, D., McAdam, A. G., & Boutin, S. (2003). Lifetime selection on heritable life-history traits in a natural population of red squirrels. *Evolution, 57*, 2416–2423.

Reale, D., & Festa-Bianchet, M. (2000). Mass-dependent reproductive strategies in wild bighorn ewes: A quantitative genetic approach. *Journal of Evolutionary Biology, 13*, 679–688.

Ritland, K., & Ritland, C. (1996). Inferences about quantitative inheritance based on natural population structure in the yellow monkey flower, *Mimulus guttatus*. *Evolution, 50*, 1074–1082.

Rodriguez, R. L., & Greenfield, M. D. (2003). Genetic variance and phenotypic plasticity in a component of female mate choice in an ultrasonic moth. *Evolution, 57*, 1304–1313.

Roff, D. A. (1995). Antagonistic and reinforcing pleiotropy: A study of differences in development time in wing dimorphic insects. *Journal of Evolutionary Biology, 8*, 405–419.

Roff, D. A., & Mousseau, T. A. (1987). Quantitative genetics and fitness: Lessons from *Drosophila*. *Heredity, 58*, 103–118.

Rolff, J., Armitage, S. A. O., & Coltman, D. W. (2005). Genetic constraints and sexual dimorphism in immune defense. *Evolution, 59*, 1844–1850.

Rønning, B., Jensen, H., Moe, B., & Bech, C. (2007). Basal metabolic rate: Heritability and genetic correlations with morphological traits in the zebra finch. *Journal of Evolutionary Biology, 20*, 1815–1822.

Routley, M. B., & Husband, B. C. (2005). Responses to selection on male-phase duration in *Chamerion angustifolium*. *Journal of Evolutionary Biology, 18*, 1050–1059.

Ryder, J. J., & Siva-Jothy, M. T. (2001). Quantitative genetics of immune function and body size in the house cricket, *Acheta domesticus*. *Journal of Evolutionary Biology, 14*, 646–653.

Sakai, A. K., Weller, S. G., Culley, T. M., Campbell, D. R., Dunbar-Wallis, A. K., & Andres, A. (2008). Sexual dimorphism and the genetic potential for evolution of sex allocation in the gynodioecious plant, *Schiedea salicaria*. *Journal of Evolutionary Biology, 21*, 18–29.

Santos, M. (1996). Apparent directional selection of body size in *Drosophila buzzatii*: Larval crowding and male mating success. *Evolution, 50*, 2530–2535.

Santos, M. (2001). Fluctuating asymmetry is nongenetically related to mating success in *Drosophila buzzatii*. *Evolution, 55*, 2248–2256.

Santos, M. (2002). Genetics of wing size asymmetry in *Drosophila buzzatii*. *Journal of Evolutionary Biology, 15*, 720–734.

Santos, M., Ruiz, A., Quezada-Diaz, J. E., Barbadilla, A., & Fontdevila, A. (1992). The evolutionary history of *Drosophila buzzatii*. XX. Positive phenotypic covariance between field adult fitness components and body size. *Journal of Evolutionary Biology, 5*, 403–422.

Sarkissian, T. S., & Harder, L. D. (2001). Direct and indirect responses to selection on pollen size in *Brassica rapa* L. *Journal of Evolutionary Biology, 14*, 456–468.

Schlichting, C. D., & Murren, C. J. (2004). Evolvability and the raw materials for adaptation. In Q. C. B. Cronk, J. Whitton, R. H. Ree, & I. E. P. Taylor (Eds.), *Plant adaptation: Molecular genetics and ecology*. Ottawa: NRC Research press.

Schluter, D. (2000). *The ecology of adaptive radiation*. Oxford: Oxford university press.

Sgro, C. M., & Hoffmann, A. A. (1998). Effects of temperature extremes on genetic variances for life history traits in *Drosophila melanogaster* as determined from parent-offspring comparisons. *Journal of Evolutionary Biology, 11*, 1–20.

Shaw, R. G., & Platenkamp, G. A. J. (1993). Quantitative genetics of response to competitors in *Nemophila menziesii*—a greenhouse study. *Evolution, 47*, 801–812.

Shaw, F. H., Shaw, R. G., Wilkinson, G. S., & Turelli, M. (1995). Changes in genetic variances and covariances, G whiz! *Evolution, 49*, 1260–1267.

Sherrard, M. E., Maherali, H., & Latta, R. G. (2009). Water stress alters the genetic architecture of functional traits associated with drought adaptation in *Avena barbata*. *Evolution, 63*, 702–715.

Simmons, L. W. (2003). The evolution of polyandry: patterns of genotypic variation in female mating frequency, male fertilization success and a test of the sexy-sperm hypothesis. *Journal of Evolutionary Biology, 16*, 624–634.

Simmons, L. W., & Garcia-Gonzalez, F. (2007). Female crickets trade offspring viability for fecundity. *Journal of Evolutionary Biology, 20*, 1617–1623.

Simmons, L. W., & Kotiaho, J. S. (2002). Evolution of ejaculates: Patterns of phenotypic and genotypic variation and condition dependence in sperm competition traits. *Evolution, 56*, 1622–1631.

Simons, A. M., Carriere, Y., & Roff, D. A. (1998). The quantitative genetics of growth in a field cricket. *Journal of Evolutionary Biology, 11*, 721–733.

Simons, A. M., & Johnston, M. O. (2006). Environmental and genetic sources of diversification in the timing of seed germination: Implications for the evolution of bet hedging. *Evolution, 60*, 2280–2292.

Simons, A. M., & Roff, D. A. (1994). The effect of environmental variability on the heritabilities of traits of a field cricket. *Evolution, 48*, 1637–1649.

Steigenga, M. J., Zwaan, B. J., Brakefield, P. M., & Fischer, K. (2005). The evolutionary genetics of egg size plasticity in a butterfly. *Journal of Evolutionary Biology, 18*, 281–289.

Stinchcombe, J. R. (2005). Measuring natural selection on proportional traits: Comparisons of three types of selection estimates for resistance and susceptibility to herbivore damage. *Evolutionary Ecology, 19*, 363–373.

Stinchcombe, J. R., Agrawal, A. F., Hohenlohe, P. A., Arnold, S. J., & Blows, M. W. (2008). Estimating nonlinear selection gradients using quadratic regression coefficients: Double or nothing? *Evolution, 62*, 2435–2440.

Stirling, D. G., Reale, D., & Roff, D. A. (2002). Selection, structure and the heritability of behaviour. *Journal of Evolutionary Biology, 15*, 277–289.

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

Theriault, V., Garant, D., Bernatchez, L., & Dodson, J. J. (2007). Heritability of life-history tactics and genetic correlation with body size in a natural population of brook charr (*Salvelinus fontinalis*). *Journal of Evolutionary Biology, 20*, 2266–2277.

Thessing, A., & Ekman, J. (1994). Selection on the genetical and environmental components of tarsal growth in juvenile willow tits (*Parus montanus*). *Journal of Evolutionary Biology, 7*, 713–726.

Thiede, D. A. (1998). Maternal inheritance and its effect on adaptive evolution: A quantitative genetic analysis of maternal effects in a natural plant population. *Evolution, 52*, 998–1015.

Thomas, M. L., & Simmons, L. W. (2008). Cuticular hydrocarbons are heritable in the cricket *Teleogryllus oceanicus*. *Journal of Evolutionary Biology, 21*, 801–806.

Tonsor, S. J., & Goodnight, C. J. (1997). Evolutionary predictability in natural populations: Do mating system and nonadditive genetic variance interact to affect heritabilities in *Plantago lanceolata*? *Evolution, 51*, 1773–1784.

Tucic, B., & Stojkovic, B. (2001). Shade avoidance syndrome in *Picea moorika* seedlings: A growth-room experiment. *Journal of Evolutionary Biology, 14*, 444–455.

Turelli, M. (1984). Heritable genetic variation via mutation-selection balance: Lerch's zeta meets the abdominal bristle. *Theoretical Population Biology, 25*, 138–193.

Van Kleunen, M., & Ritland, K. (2004). Predicting evolution of floral traits associated with mating system in a natural plant population. *Journal of Evolutionary Biology, 17*, 1389–1399.

Van Tienderen, P. M. (2000). Elasticities and the link between demographic and evolutionary dynamics. *Ecology, 81*, 666–679.

Verhoeven, K. J. F., Biere, A., Nevo, E., & Van Damme, J. M. M. (2004). Differential selection of growth rate-related traits in wild barley, *Hordeum spontaneum*, in contrasting greenhouse nutrient environments. *Journal of Evolutionary Biology, 17*, 184–196.

Wagner, G. P. (1989). Multivariate mutation-selection balance with constrained pleiotropic effects. *Genetics, 122*, 223–234.

Wagner, G. P., & Altenberg, L. (1996). Complex adaptations and evolution of evolvability. *Evolution, 50*, 967–976.

Wagner, G. P., Booth, G., & Bagheri-Chaichian, H. (1997). A population genetic theory of canalization. *Evolution, 51*, 329–347.

Wagner, G. P., Pavlicev, M., & Cheverud, J. M. (2007). The road to modularity. *Nature Reviews Genetics, 8*, 921–931.

Walsh, B., & Blows, M. W. (2009). Abundant genetic variation + strong selection = multivariate genetic constraints: a

geometric view of adaptation. *Annual Review of Ecology, Evolution and Systematics, 40*, 41–59.

Watkins, T. B. (2001). A quantitative genetic test of adaptive decoupling across metamorphosis for locomotor and life-history traits in the Pacific tree frog, *Hyla regilla*. *Evolution, 55*, 1668–1677.

Wayne, M. L., Hackett, J. B., & Mackay, T. F. C. (1997). Quantitative genetics of ovariole number in *Drosophila melanogaster*. 1. Segregating variation and fitness. *Evolution, 51*, 1156–1163.

Weber, S. L., & Scheiner, S. M. (1992). The genetics of phenotypic plasticity. IV. Chromosomal localization. *Journal of Evolutionary Biology, 5*, 109–120.

Weller, S. G., Sakai, A. K., Culley, T. M., Campbell, D. R., & Dunbar-Wallis, A. K. (2006). Predicting the pathway to wind pollination: heritabilities and genetic correlations of inflorescence traits associated with wind pollination in *Schiedea salciaria* (Caryophllaceae). *Journal of Evolutionary Biology, 19*, 331–342.

Willi, Y., van Buskirk, J., & Hoffmann, A. A. (2006). Limits to the adaptive potential of small populations. *Annual Review of Ecology, Evolution and Systematics, 37*, 433–458.

Wilson, A. J. (2008). Why h^2 does not always equal V_A/V_P*? Journal of Evolutionary Biology, 21*, 647–650.

Wilson, A. J., Coltman, D. W., Pemberton, J. M., Overall, A. D. J., Byrne, K. A., & Kruuk, L. E. B. (2005). Maternal genetic effects set the potential for evolution in a free-living vertebrate population. *Journal of Evolutionary Biology, 18*, 405–414.

Wilson, A. J., Hutchings, J. A., & Ferguson, M. M. (2003). Selective and genetic constraints on the evolution of body size in a stream-dwelling salmonid fish. *Journal of Evolutionary Biology, 16*, 584–594.

Windig, J. J. (1994). Reaction norms and the genetic basis of phenotypic plasticity in the wing patern of the butterfly *Bicyclus anynana*. *Journal of Evolutionary Biology, 7*, 665–695.

Winn, A. A. (2004). Natural selection, evolvability and bias due to environmental covariance in the field in an annual plant. *Journal of Evolutionary Biology, 17*, 1073–1083.

Woods, R. E., Hercus, M. J., & Hoffmann, A. A. (1998). Estimating the heritability of fluctuating asymmetry in field *Drosophila*. *Evolution, 52*, 816–824.

Worley, A. C., & Barrett, S. C. H. (2000). Evolution of floral display in *Eichhornia paniculata* (Pontederiaceae): genetic correlation between flower size and number. *Journal of Evolutionary Biology, 14*, 469–481.