Title: Evolutionary rates for multivariate traits: the role of selection and genetic variation.

Authors: William Pitchers\textsuperscript{1,3}, Jason B. Wolf\textsuperscript{2}, Tom Tregenza\textsuperscript{3}, John Hunt\textsuperscript{3*} and Ian Dworkin\textsuperscript{1*}

Affiliations:

\textsuperscript{1}Department of Zoology, Program in Ecology Evolutionary Biology and Behavior, BEACON Center for the study of Evolution in Action, Michigan State University, East Lansing, MI 48824, USA.

\textsuperscript{2}Department of Biology & Biochemistry, University of Bath, Claverton Down, Bath, BA2 7AY, UK.

\textsuperscript{3}College of Life and Environmental Sciences, Centre for Ecology and Evolution, University of Exeter, Tremough Campus, Penryn, Cornwall, TR10 9EZ, UK

*Corresponding Authors: John Hunt – J.Hunt@exeter.ac.uk & Ian Dworkin

idworkin@msu.edu

Keywords: Quantitative genetics, evolution, haldanes, G matrix, natural selection, evolutionary rates, constraint, evolvability, scaling.

Supplemental Material: Data & scripts available at DRYAD DOI:xxxxx
Summary

A fundamental question in evolutionary biology is the relative importance of selection and genetic architecture in determining evolutionary rates. Adaptive evolution can be described by the multivariate breeders' equation \( \Delta \bar{z} = G \beta \), which predicts evolutionary change for a suite of phenotypic traits \( \Delta \bar{z} \) as a product of directional selection acting on them \( \beta \) and the genetic variance-covariance matrix for those traits \( G \). Despite being empirically challenging to estimate, there are enough published estimates of \( G \) and \( \beta \) to allow for synthesis of general patterns across species. We use published estimates to test the hypotheses that there are systematic differences in the rate of evolution among trait types, and that these differences are in part due to genetic architecture. We find evidence that sexually selected traits exhibit faster rates of evolution compared to life-history or morphological traits. This difference does not appear to be related to stronger selection on sexually selected traits. Using numerous proposed approaches to quantifying the shape, size and structure of \( G \) we examine how these parameters relate to one another, and how they vary among taxonomic and trait groupings. Despite considerable variation, they do not explain the observed differences in evolutionary rates.
Introduction

Predicting the rate and direction of phenotypic evolution remains a fundamental challenge in evolutionary quantitative genetics [1-4]. Empirical studies have demonstrated that most traits are heritable [5-8] and can respond to selection – a prediction borne out by an abundance of short- (e.g. [9-11]) and long-term (e.g. [9,12-14]) artificial selection experiments targeting single traits. However, in most biological systems, the targets of selection are suites of traits. Furthermore, different traits are tied together by genetic associations (typically quantified as covariances), and consequently selection on one trait can lead to evolutionary changes in other traits [7,8,11,15-21]. Indeed, genetic covariation between traits appears to be ubiquitous and has the potential to shape the evolution of associated traits [7,10,17,18,20,22,23]. Therefore, to improve our understanding of phenotypic evolution it is necessary to invoke a multivariate perspective [5,17-19,24].

The evolutionary response of a suite of traits can be predicted by the multivariate breeder’s equation $\Delta \mathbf{X} = \mathbf{G}\beta$ where $\Delta \mathbf{X}$ is the vector of responses in phenotypic means for the suite of traits under consideration, $\mathbf{G}$ is the additive genetic variance-covariance matrix and $\beta$ is the vector of linear (directional) selection gradients [5-8]. The importance of $\mathbf{G}$ to phenotypic evolution can be illustrated using the concept of “genetic degrees of freedom” [9,11,15]. Whenever there is genetic covariance between them, the number of trait “combinations” in $\mathbf{G}$ that can respond to selection can be considerably smaller than the actual number of measured traits. This can be true even
when each trait in $G$ is heritable and all pairwise genetic correlations between them are less than one [1-3,9,11,25]. This reduced dimensionality constrains the population to evolve in a genetic space with fewer dimensions than the number of traits (and trait combinations) potentially under selection. A matrix whose variance is concentrated in one or a few dimensions can exhibit “lines of least evolutionary resistance”; directions in which the multivariate evolutionary response can proceed more rapidly than in others [15]. The presence of these lines of least evolutionary resistance can have a major influence in biasing the direction of evolutionary trajectories (Figure 1; ref.s [7,11,15-20]), making the $G$ matrix more informative about the short term capacity of a population to respond to selection (i.e. its evolvability) than the heritabilities of individual traits [7,10,17,18,20,22,23].

A variety of measures have been proposed as proxies for the evolutionary potential of a population. Most current approaches represent a function of the components of the multivariate breeder’s equation: $G$, $\beta$ and $\Delta\bar{F}$ [5,17-19,21,24]. Unfortunately, few studies simultaneously estimate more than one of these components. The few notable exceptions suggest that the structure of $G$ plays an important role in directing phenotypic evolution [26-29]. Even fewer studies provide direct estimates of observed rates of evolution [30,31]. However, many individual estimates of selection and evolutionary rates exist in the literature and evolutionary research has benefitted from reviews that synthesize these parameters [30-38]. There is considerable variation in the strength of selection across different trait types and fitness measures [33,34,38], as well
as over time (but see ref.s [36,39,40]. On average, linear selection appears stronger on
morphological than life-history traits and both linear and quadratic selection is stronger
when acting on mating success and fecundity compared to viability [1-4,33,38]. It should
be acknowledged, however, that such studies are subject to some methodological
debate [5-8,35] and potentially publication biases [9-11,40]. In particular, there has
been considerable disagreement about trait scaling, and how it influences estimates and
broader evolutionary conclusions [19,22,41].

Although they have not received the same attention as selection gradients, reviews
based on published genetic parameters also show clear differences across trait types.
Morphological traits generally have higher heritabilities than life-history traits, with
physiological and behavioural traits intermediate between these extremes [9,12-14,32],
but see [6-8,11,15-21]. Sexual traits have also been shown to have higher additive
genetic variances compared to non-sexually selected traits [7,10,17,18,20,22,23,42],
although this finding is based on few studies. As discussed above, trait scaling has been
shown to alter the observed patterns [19,22,41].

There have been even fewer attempts at synthesis from a multivariate perspective.
Notably, Kirkpatrick [20], Kirkpatrick & Lofsvold [9], Agrawal & Stinchcombe [23], and
Schluter [11,15] collected small samples of G matrices from the literature and found
that much of the available variance was concentrated in the first few dimensions. This
suggests that few genetic degrees of freedom may be the norm, but we know of no
systematic review that reveals how general this pattern is or whether it differs across taxa or trait types. Likewise, although reviews on the rate of contemporary microevolution suggest that rapid evolution should be viewed as the norm rather than the exception [15,30,31], a comprehensive review of evolutionary rates across different taxa and trait types does not currently exist.

We compiled a database of reported genetic parameters from the literature to ask whether different types of traits evolve at different rates, and whether such differences correlate with differences in selection, in patterns of genetic (co)variation or both. We performed a quantitative literature review, to examine whether observed rates of evolutionary response differ across trait types (morphological, life-history and sexual) in plants and animals. We relate these observed rates of evolutionary response to estimates of linear and quadratic selection, as well as measures that capture the size, shape and structure of $G$ [7,11,15-20], to determine whether there is an association across trait types and taxa. We find evidence that sexual traits evolve faster than other traits in animals but not in plants, where life-history traits evolve fastest. These increased rates of evolution do not appear to be attributable to the same cause however. In plants we find that selection also appears to be strongest on life-history traits, whereas in animals selection on sexually selected traits appears to be stronger than on life-history but indistinguishable from that on morphology. We then examined how the measures used to capture the size, shape and structure of $G$ vary among trait types and between taxa, but find that this incompletely explains the observed pattern of
evolutionary rates. In addition, we compare the various measures based upon $G$, and show that for these empirically observed matrices, many strongly co-vary.

Methods

Compilation of Database

We compiled our datasets by searching for publications on the ISI Web of Science database between March 2006 and August 2012. We then refined this preliminary list of references on the basis of their title, abstract and keywords and attempted to obtain the full text for all papers included in the dataset.

Rates of evolution have been measured using a number of different units, most prominently darwins [7,10,17,18,20,22,23,43,44] and haldanes [5,17-19,21,24,43,45]. Measurements in darwins have proved most appropriate for researchers studying evolution on macro-evolutionary scales (e.g. paleontologists), since they express the rate of evolution per million years (although there are known methodological issues with making comparisons [44,46]). However, for our purposes rates expressed in haldanes are the appropriate unit as they measure change per generation and are used to measure evolution on a micro-evolutionary scale – the scale over which $G$ may be important. We therefore compiled a database of evolutionary rate measured in haldanes only. We performed searches for the terms ‘rate of evolution’, ‘rate of adaptation’, ‘haldanes’, ‘response to selection’ and ‘experimental evolution’. This process was aided considerably by making use of the measurements from the studies
previously compiled by Hendry et al [26-29,47]. Where studies reported the results of experimental evolution without explicitly reporting a rate of response, we contacted the authors to ask for the data needed (e.g. generation time) to calculate a rate in haldanes, standardizing traits as necessary. Previous work has shown that even with log transformation of ratio scale data (where means and variances might co-vary), this had little influence on overall estimates for haldanes [31].

For the database of selection gradients, we began with the database compiled by Kingsolver et al [30,31,33,37], and supplemented this with additional measures from work published after 2001 by searching for the terms ‘natural selection’, ‘sexual selection’, ‘selection gradient’ or ‘selection differential’. Unlike Kingsolver et al [30-38] we included both field and laboratory studies. While there has been discussion about the effects of trait scaling (mean vs. standard deviation) on estimates of selection [19,35], we have only included estimates standardized using the approach as advocated by Lande and Arnold [21], which has been broadly used.

For the G matrix dataset we searched the Web of Science database using the terms ‘G matrix’ (or ‘G-matrix’), ‘covariance matrix’ (or ‘co-variance matrix’ or ‘(co)variance matrix’) or ‘quantitative genetics’. We recorded G matrices expressed both as genetic (co)variances (provided we were able to mean-standardize them, following [19]) and as genetic correlations and narrow sense heritabilities. Where possible (i.e. where estimates of phenotypic variance had been presented alongside genetic correlations and heritabilities) we back-calculated the genetic variances and covariances as: $V_A = h^2 V_p$ and
where $V_A$ and $V_P$ are the additive genetic and phenotypic variances,

$h^2$ is the narrow sense heritability and $r_G$ is the genetic correlation between traits $x$ and $y$. In cases where matrices were incomplete we contacted the author(s) to request the missing estimates. We thus have two $G$ datasets; correlation matrices and covariance matrices. Since we found correlations to be reported more often than covariances, the correlation dataset is a superset of matrices that includes those in the covariance dataset. Trait scaling for the co-variance matrices is discussed below. In a number of cases matrices had component traits that had been measured in difficult-to-compare units (e.g. both a length and a volume), or where traits were expressed as residuals (e.g. from regression against size). In these cases we excluded these from the reported analysis, but inclusion had little effect on the results. A number of matrices were also found to include cells with correlations $>1$ and in these cases we rounded the offending correlation down to 1.

Defining Trait Categories and Measures

Since we wished to make comparisons across different ‘trait types’ (sensu [33,34,38]), it was necessary to assign our measurements from the literature into categories. We chose three trait categories: life-history, morphological and sexually selected traits. It is relatively straightforward to separate life-history from morphological traits and the majority of measurements in the literature fall into these two categories. In animals, we defined sexual traits as those where we were able to find at least one study demonstrating the trait was subject to female preference or used in male-male competition. For plants, we defined floral morphology as sexually selected
[36,39,40,48]. Thus, for both plants and animals, our sexually selected and morphology categories are not mutually exclusive. In an attempt to reduce error in our study, traits that did not fit clearly into one of our three categories were excluded from our dataset. For $G$ matrices whose component traits did not all fit the same category, we split the matrix to produce sub-matrices relating to traits only within a single category. Where matrices contained a single trait whose category differed from all others in the matrix we removed that trait from the matrix.

When making comparisons across our trait categories, we acknowledge that our classifications may not be directly equivalent in plants and animals. We therefore included a ‘taxon’ category in our statistical models. The list of individual measures of evolutionary rate was treated as a single response variable, as were the standardized selection gradients.

In our analysis of the $G$ data, we wished to capture those attributes of $G$ that might be expected to influence the rate of evolutionary change. Matrices vary principally in terms of size and structure. While numerous studies suggest that the alignment of axes of $G$ with $\beta$ is likely to be important, the nature of the data we were able to compile does not allow us to quantify alignment. Instead (as outlined below) we utilized a number of scalar measures derived from $G$, meant to capture aspects of the size and structure as a means to express evolutionary potential, All of the measures we used are summarized in Table 1. One general concern is that not all of the measured we used explicitly accounted for the number of traits included in the matrix (i.e. $n_D$). While, in general the
number of traits seemed to have a small influence on these measures (Figures 4 & 5),

we also took several steps to account for these effects, such as including number of

traits as a co-variante in the models (below) and also by examining the effects of scaling

$n_D$ by either trait number or its square (“effective subspace”, as suggested by one of the

manuscript referees). In none of these cases did it substantially alter the results. While

we use the name “effective dimensionality” for $n_D$, as proposed by Kirkpatrick [20], this

measure actually captures aspects of matrix eccentricity, not dimensionality.

For the dataset of $G$ as mean-standardized covariance matrices we used the three $G$-

structure measures suggested by Kirkpatrick [20]: ‘total genetic variance’ ($tgv$),

‘maximum evolvability’ ($e_{\text{max}}$) & ‘effective number of dimensions’ ($n_D$), and also Hansen

and Houle’s [19] ‘average evolvability’ ($\bar{e}$). For the dataset of correlation matrices, we

calculated Pavlicev et al.’s [49] eigenvalue variance ($\text{Var}(\lambda)$) and relative eigenvalue

variance ($\text{Var}_{\text{rel}}(\lambda)$) and also Agrawal & Stinchcombe’s [23] eigenvalue evenness ($E$).

Both sets of $G$ matrix measures are defined in Table 1.

While we present results from analyses of both the (co)variance and correlation matrix

datasets, it is important to note that results are not directly comparable between them,

since it is well known that different methods of scaling (i.e. mean-standardizing

(co)variance matrices vs. effectively variance-standardized correlation matrices)

produce fundamentally different results for genetic attributes [6,19,35]. Furthermore,

though the correlation matrix dataset is larger, we note that the covariance – not
correlation – matrix is part of the fundamental formulation for response to selection [21].

Statistical Analyses

Analyses were performed using R (version 2.13.0; ref. [50]); we fit generalized linear mixed-effect models using the MCMCglmm package (version 2.15; ref. [51]). A large proportion of studies reporting selection gradients also reported standard errors or confidence intervals (from which standard errors can be calculated). As noted by Kingsolver et al [38], this allows for the application of formal meta-analyses, and we followed their lead in modelling selection data with a meta-analysis including random-effects to account for study- and species-level autocorrelation. We analysed estimates of standardised selection gradients ($\beta$) expressed as absolute values.

We found that standard errors or confidence intervals were reported much less frequently among studies of G or rates of evolution, and so we were unable to account for measurement error variance in these analyses as we had for selection, though the model structure we used was otherwise similar. We fit a set of models, and then evaluated model fit by comparing Deviance Information Criterion values (DIC) [52], and confirmed our selections by refitting the model set using reduced maximum likelihood (lme4 package [53]) and comparing fits using Bayesian Information Criterion scores (BIC) and likelihood ratio tests (Supplemental Methods) using a parametric bootstrap. The selected models for each dataset are described in Table 2, and full model sets are available with the data and scripts on Dryad. Since we modelled the magnitude
(absolute value) of our response variables, the appropriate distribution is the folded normal [38]. We therefore extracted the posterior distributions of solutions, took the mean and standard deviation from these distributions and applied these to the folded normal distribution. We then report the mean and credible intervals from these corrected distributions [38].

Our data contained 2571 estimates of the rate of evolutionary response (measured in haldanes); there were comparatively few estimates for plants, with no estimates available on the observed rate of evolution for sexually selected (floral) traits. This imbalance caused our estimates to be unstable so we modelled plant and animal rates separately. We had 776 estimates of $\beta$, but $G$ is reported less frequently in the literature (Table 3) and our sample size of $G$ measures was 81 (co)variance matrices and 221 correlation matrices.

**Results**

Observed rates of evolution differ among trait types and between plants and animals

The overall posterior mean for evolutionary rate was 0.13 haldanes, with a 95% credible interval from 0.08 – 0.17. Credible intervals for estimates in plants are quite wide (Figure 2), most likely due to the comparatively low number of studies in these categories. However there is a clear trend for faster rates in life-history traits, with the life-history estimate being 2.02 times as large (95% credible interval 1.00 – 3.03) as that for morphology (Table 3). In animals, life-history and morphology have similar
estimates, but the posterior mean estimate for sexually selected traits is higher – 1.48
times that for morphology (95% CI 0.78 – 2.11), and 1.51 times that for life-history (95%  
CI 0.50 – 2.54). These results are consistent with similar rates of evolution for  
morphology in both plants and animals, with higher rates for life-history traits in plants  
and for sexually selected traits in animals.

*Standardised selection gradients show different patterns between plants and animals*

The overall posterior mean for absolute linear selection gradients was 0.21 (95% CI =  
0.17 – 0.26), which was somewhat higher than the estimate reported by Kingsolver et al.  
[38] (0.14, 95% CI = 0.13 – 0.16), most likely due to our inclusion of lab studies. The  
credible intervals from our full model are again wider for plants, likely reflecting smaller  
sample size (Table 3). For both plants and animals there is little difference between the  
estimates for morphological and sexually selected traits. In plants, the model suggests  
that selection is stronger on life-history traits, whose estimate is 40% larger than that for  
morphology and approximately twice that for sexually selected traits. By contrast, in  
animals selection appears to be weaker for life-history; the estimate for selection on  
life-history traits is 0.43 times (95% CI 0.11 – 0.97) that for morphology, and 0.49 times  
(95% CI 0.17 – 0.80) that for sexually selected traits (Figure 3).

*The marginal utility of multiple measures*

The magnitude, shape and alignment of the G matrix all have the potential to influence  
the rate of evolution, but with the data available we are able to use measures intended  
to quantify only the first two of these properties. Of the measures we report $tgv, e_{\text{max}}$
and $\bar{e}$ can be thought of as measures of magnitude, whereas $n_D$, $\text{Var}(\lambda)$, $\text{Var}_{rel}(\lambda)$ and $E$.

are intended to quantify the departure of the matrix from symmetricality (how
dissimilar variances are along the multiple axes of $G$). It is immediately obvious that the
magnitude measures are doing a good job of quantifying the same property of each
matrix (Table 1, Figures 4 & 5), since $tgv$, $e_{\text{max}}$ and $\bar{e}$ are all very strongly inter-correlated
($r > 0.96$ in all cases). Given that these measures of magnitude are also strongly
correlated ($r > 0.93$ in all cases) with the magnitude of $g_{\text{max}}$ (i.e. the principal eigenvalue
of $G$), it is perhaps unsurprising that they are only poorly predicted by the number of
traits measured, with which they are correlated only at $r = 0.15 - 0.19$.

With respect to the measures of matrix eccentricity, the first thing we note is that $\text{Var}(\lambda)$
and $\text{Var}_{rel}(\lambda)$ are strongly correlated with each other ($r = 0.87$), and negatively correlated
with $E$. ($r = -0.32$ & -0.55 respectively). Though $E$ was defined as a measure of
correlation matrices [23], when we applied the evenness formula to our dataset of
covariance matrices we find that the resulting measure is strongly correlated with
Kirkpatrick’s [20] $n_D$ ($r = 0.82$).

The structure of $G$

We performed separate analyses and model selection procedures for each of our
measures describing the structure of $G$. Our models comparing covariance matrices
revealed very similar patterns of estimates for $e_{\text{max}}$, $tgv$ and $\bar{e}$. Furthermore the pattern
of estimates among trait types was consistent between plants and animals (Figure 6). In
all cases the estimates for life-history and sexually selected traits were similar and those
for morphology were higher, but with much overlap in credible intervals our confidence
in these differences is low. Our results for $n_D$ also show consistent patterns of estimates
between plants and animals, with the estimates showing a shallow increasing trend
from life-history to morphology to sexually selected traits (Figure 6(d)), but once again
there is wide overlap among credible intervals, indicating low confidence in this trend.

The results of our analyses of $G$ matrices expressed as correlations were more diverse.
The pattern of estimates for $\text{Var}(\lambda)$ showed a trend for values to increase from life-
history to morphology to sexually selected traits in both plants and animals, though the
estimates for animals were larger than those for plants (Figure 7(a)). A similar trend was
present in estimates for $\text{Var}_{\text{rel}}(\lambda)$ for animals, thought the direction of the trend is
reversed in plants (Figure 7(b)). The wide overlap of credible intervals indicates low
confidence in both these trends however. Finally, our estimates for $E$ show a decreasing
trend from life-history to morphology to sexually selected traits in both plants and
animals, with higher estimates for plants than for animals (Figure 7(c)).

**Discussion**

Predicting the rate and direction of phenotypic evolution remains a fundamental
challenge in evolutionary genetics [1-4,54], with the multivariate breeders’ equation
playing a central role. Researchers have published estimates of selection and of $G$ from
many systems (though not commonly both in the same system). Here we have
integrated these data to ask if some traits generally evolve more rapidly than others,
and whether any differences correspond to differences in selection, $G$ or both. Reviews
such as this are unavoidably limited by the availability of published genetic parameters, and the resulting imbalances in the data. Nevertheless, we find that in animals, but not plants, sexual traits evolve ~50% faster than morphological and life-history traits. We find no evidence that the increased rate of evolution observed for sexual traits was due to stronger selection operating on these traits relative to morphological and life-history traits. We found weak evidence for differences in the evolutionary potential of $G$ among trait types, though this fails to provide a satisfactory explanation for the increased rates of evolution associated with sexually selected traits. In addition, we confronted our compiled data sets with a number of the scalar measures used to examine the size, shape and structure of $G$ (Table 1). We observed that many of these measures have considerable shared information (Figures 4 & 5), though in general this pattern of relationships would recommend the use of at least two measures; one to express the magnitude of $G$ and a second relating to evenness/variance of the eigenvalues. While there may be particular instances where these measures result in widely divergent estimates, at least with respect to the empirical estimates we have collated, the marginal benefits of using all of them are an illustration of diminishing returns. We speculate that one potential use (which would require considerable additional research) would be in a fashion analogous to the molecular population geneticists’ use of the parameter Tajima’s $D$, which is simply a scaled measure of two different estimates of the population mutation rate, $4N_e \mu$. It is possible that subtle differences among scalar $G$ measures may ultimately provide important insights into the structure of $G$. One surprising observation that emerges from our results, is that the number of traits used
to estimate $G$ is not well correlated with any of the scalar measures we used. One explanation for this becomes clear when considering that the magnitudes of the principal eigenvalue of $G$ is so highly correlated with total genetic variation (the trace of $G$). This suggests that an overwhelming proportion of all of the variation is found along this principal vector (which would differ for each $G$). However, it is well known that estimating $G$ can be difficult and insufficient sampling at the level of families can inflate the magnitude of the principal eigenvalue [55,56].

Reviews based on published estimates of evolutionary rates [30,31] have provided a number of important insights into the evolutionary process. Hendry & Kinnison [30] provided the theoretical foundations for measuring evolutionary rates and used a small sample of published estimates to propose that rapid evolution should be viewed as the norm rather than the exception. In a much larger study, however, Kinnison & Hendry [31] showed that the frequency distribution of evolutionary rates measured in haldanes is log-normal (i.e. many slow rates and few fast rates, median haldanes = $5.8 \times 10^{-3}$) and that life-history and morphological traits appear to evolve equally as fast when measured in haldanes. In agreement with these reviews, we found that the frequency distribution of evolutionary rates in our study was also log-normal and that the median rate across trait types and taxa was similar (median haldanes = $7.6 \times 10^{-3}$) to that reported in Kinninson & Hendry [31]. Moreover, we found little evidence to suggest that the evolutionary rates of life-history and morphological traits differed in either plants or animals. A novel outcome of our analysis, however, is the finding that sexual traits evolve faster than life-history and morphological traits in animals. Our findings provide
evidence for a general pattern of faster evolution in sexual traits in animals to add to the
highly cited individual examples of very rapid evolution of sexual traits [57,58] and their
role in speciation [59,60]. More work is needed to determine whether this pattern also
exists in plants.

Reviews synthesizing estimates of selection gradients are far more extensive in the
literature [33-39]. In their seminal review, Kingsolver et al. [33] found that the frequency
distributions of absolute linear and quadratic selection gradients were exponential and
generally symmetrical around zero, suggesting that stabilizing and disruptive selection
occur equally frequently and with similar strength in nature. Kingsolver et al. [33] also
found that linear selection was stronger on morphological than life-history traits. The
most recent review [38] containing an updated data set and using formal Bayesian
meta-analysis to control for potential biases [34,35,37] confirmed many of the main
findings of Kingsolver et al. [33], with the notable exception that linear selection appears
stronger in plants than animals.

In agreement with this most recent review [38], we found that the distribution of
absolute linear and quadratic selection gradients were exponential (see Supplement).
However, our estimates for absolute linear selection gradients were higher than
reported by Kingsolver et al. [38] (0.24 (0.17, 0.26) versus 0.14 (0.13, 0.16)). This raises
the obvious question of why this difference occurs. There has been much discussion on
the general limitations of using selection gradients in synthetic reviews (e.g.
[33,35,37,38] and these arguments undoubtedly also apply to our study. However, as
most of these limitations are inherent to both studies, it is unlikely that they explain the observed differences. Furthermore, given that we used the same Bayesian framework as Kingsolver et al. [38] it is also unlikely that our analytical approach has generated the observed differences. The most likely reason for the observed differences is the way that traits and taxa were categorized across these studies. Kingsolver et al. [38] used four different trait categories (size, morphological (not including size), phenology and life-history (not including phenology)) and categorized taxa as invertebrates, vertebrates or plants in their analysis. In contrast, we only distinguished between animals and plants and used three different trait categories (morphological, life-history and sexual) in our analysis, the latter of which includes a mixture of morphological and behavioural traits. Thus, there are likely to be some differences in how selection gradients are distributed amongst categories in our analyses compared to those in Kingsolver et al. [38] and this may account for the observed differences between studies. Irrespective of the underlying reasons for these differences, our main finding that there is little difference in selection gradients across trait types and taxa suggests that selection alone is unlikely to explain the higher rate of evolutionary response we observe for sexual traits in animals.

After decades of quantitative genetic research it is now widely accepted that the additive genetic variance-covariance matrix (G) plays a major role in facilitating/constraining phenotypic evolution [16,19,20]. The way in which G shapes phenotypic evolution can be envisaged using the concept of genetic degrees of freedom (Figure 1; [9,15]). Whenever there is genetic covariation between the individual traits
contained in $G$, there is the potential for fewer axes of genetic variation than observed traits [9,15,61,62] (but see [63]), which can influence evolutionary rates [64]. Where the majority of the genetic variance is concentrated in a few direction – known as “lines of least evolutionary resistance” [15] – these have been shown to play an important role in directing the short-term evolutionary trajectory of a population [15,65-69]. Quantifying these properties of $G$ is an essential step if we are to explore these ideas empirically. Our comparison of a number of the measures that have been proposed to fit this role indicates that, however informative it may or may not prove to be, we are able to reliably quantify the magnitude of $G$, since we find broad agreement among magnitude measures. Perhaps unsurprisingly, it seems that the magnitude of a matrix is somewhat more straightforward to describe with a scalar measure than the eigenvalue evenness/eccentricity/dimensionality. The measures available for quantification of the shape of $G$ in multiple dimensions are much less tightly inter-correlated than those dealing with matrix magnitude when compared using empirical data. What this ultimately means for our understanding of evolvability is unclear, but it is important to acknowledge the gaps in our current understanding if we are to progress.

The finding that genetic variance for sexual traits may be spread less evenly across dimensions in animals runs counter to our hypothesis, and suggests that the potential for genetic constraints does not explain the higher rate of evolution we observe for these traits. We found at best, only weak evidence for differences in the measures to capture the size and shape of $G$ with respect to our trait groupings. There has been debate over the importance of sexual selection in plants [70], but there is theoretical
and empirical [71] evidence suggesting that floral morphology is indeed subject to sexual selection. Unfortunately though, there are currently no data on evolutionary rates for sexual traits in plants, making it difficult to understand the implications of this increased dimensionality. Our findings indicate that the subject warrants greater attention. Additionally, when considering these issues, researchers also need to keep in mind that their decisions about measurement scaling issues are likely to be important when measuring selection [35] and genetic variability [6]. This is especially important when addressing the question of evolvability, where both these measures must be brought together [19].

Collectively, our results suggest that the higher rate of evolution observed for sexual traits in animals is only weakly associated with these scalar measures summarizing $G$ for these traits rather than stronger selection. However, as our data set is based on derived estimates of evolutionary rates, standardized selection gradients and $G$, there are a number of inevitable limitations that apply to our findings. First, there are limitations with using the matrix structure measures ($n_D$, $E\lambda$, $Var(\lambda)$ or $Var_{rel}(\lambda)$) to capture the dimensionality of $G$ [20]. Although these measure are calculable from published estimates of $G$, they do not explicitly test how many of the dimensions of $G$ actually exist (i.e. have statistical support). A number of approaches [63,72] have been taken to directly estimate the dimensionality of $G$ [72,73] though such studies have found both populations that have evolutionary access to all dimensions of $G$ [63] and others that are constrained by LLER’s [72,74]. Second, our analysis does not consider the alignment between the vectors of selection and $G$. LLER’s only constrain the response to selection
when they are poorly aligned with vectors of selection [26,28,75]. These limitations can only be resolved by further analysis of the raw data sets contained in the original studies we review. This is particularly true for better estimation of $G$ itself, as well as its actual dimensionality, which can only be performed with the raw data [56,61,63,76-79]. Future studies would therefore greatly benefit from researchers publishing their raw datasets in open repositories [80] and we strongly encourage researchers to do so. Our database can be found at DRYAD DOI:xxxxxxxx.

471

Acknowledgements

472 We thank S.F. McDaniel and W.U. Blanckenhorn for donating unpublished matrices and to R. Snook, S. Chenoweth, T. Chapman, R. Firman, and M. Reuter for data on experimental evolution. I. Stott, F. Ingleby and two anonymous reviewers provided comments on an earlier draft of this manuscript. JH, TT and JBW were funded by NERC, JH and JBW by the BBSRC and JH by a University Royal Society Fellowship. WP was supported by a NERC studentship (awarded to JH and TT) and ID & WP were funded by NIH grant no. 1R01GM094424.

480

References


484 2. Merilä, J., Kruuk, L. & Sheldon, B. 2001 Natural selection on the genetical


26

171, 972–986. (doi:10.1086/656512)


31. Kinnison, M. & Hendry, A. 2001 The pace of modern life II: from rates of contemporary microevolution to pattern and process. Genetica 112, 145–164.


53. Bates, D., Maechler, M. & Bolker, B. In press. lme4: Linear mixed-effects models using S4 classes.


67. Renaud, S., Auffray, J.-C. & Michaux, J. 2006 Conserved phenotypic variation patterns, evolution along lines of least resistance, and departure due to


**Figure Legends**

**Figure 1.** The effect of $g_{\text{max}}$ on the response to selection where traits genetically covary.

The axes represent the breeding values for 2 hypothetical traits. The population mean is at the solid point and the surrounding ellipse is the 95% confidence region for the distribution of trait values about the mean. That these traits covary is evident as the ellipse is at an angle relative to the trait axes. The axes of the ellipse represent the 2 orthogonal directions (eigenvectors) of variance present – there is more standing genetic variance along the major axis ($g_{\text{max}}$) than the minor axis. They grey lines are ‘contours’ on a fitness landscape, with an adaptive peak at ‘S’. Rather than evolving directly toward the peak (dashed arrow), the influence of $g_{\text{max}}$ may cause the population to evolve along an indirect course (bold arrow). In some cases this may even, result in the population evolving toward an alternate fitness peak (e.g. at ‘A’, modified contours not shown) in line with $g_{\text{max}}$, even though it is more distant from the current mean.

**Figure 2.** Posterior means and 95% credible intervals for estimates of absolute rate of evolution (haldanes). Open points are for plants and filled points for animals. Trait types are life-history (LH), morphology (M) and sexually selected (S) and filled points are for animals and open points for plants (no data available for sexual traits in plants).
Figure 3. Posterior means and 95% credible intervals for estimates of standardized selection gradients ($\beta$) by trait type. Trait labels and taxon symbols are as in Figure 2.

Figure 4. Pairs plot to illustrate the relationships between measures used to describe the structure of G expressed as covariance matrices. Measures are ‘total genetic variance’ ($tgv$), ‘maximum evolvability’ ($e_{max}$) & ‘effective number of dimensions’ ($n_D$) [20], the first eigenvalue of G ($g_{max}$), ‘average evolvability’ ($\bar{e}$) [19], ‘eigenvalue evenness’ ($E\lambda$ – originally intended for use with correlation matrices [23]) and the number of traits included in the matrix ($k$). Figures in the lower off-diagonal are pairwise correlations between the measures.

Figure 5. Pairs plot to illustrate the relationships between measures used to describe the structure of G expressed as correlation matrices. Measures are ‘eigenvalue variance’ (Var($\lambda$)) [49], ‘eigenvalue evenness’ ($E\lambda$) [23], ‘relative eigenvalue variance’ ($\text{Var}_{\text{rel}}(\lambda)$) [49], the first eigenvalue of G ($g_{max}$) and the number of traits included in the matrix ($k$). Figures in the lower off-diagonal are pairwise correlations between the measures.

Figure 6. Posterior means and 95% credible intervals for the four measures used to characterise G matrices expressed as covariances (see methods section); (a) ‘maximum evolvability’ ($e_{max}$), (b) ‘total genetic variance’ ($tgv$), (c) ‘average evolvability’ ($\bar{e}$) and (d)
‘effective dimensionality’ \( (n_D) \). Trait types are life-history (LH), morphology (M) and sexually selected (S) and filled points are for animals and open points for plants.

**Figure 7.** Posterior means and 95% credible intervals for the four measures used to characterise \( G \) matrices expressed as correlations (see methods section); (a) ‘eigenvalue variance’ \( \text{Var}(\lambda) \), (b) relative eigenvalue variance’ \( \text{Var}_{\text{rel}}(\lambda) \) and (c) ‘eigenvalue evenness’ \( E_\lambda \). Trait labels and taxon symbols are as in Figure 6.
Table 1. G-matrix metrics used in this study. Eigenvalue variance, relative eigenvalue variance and eigenvalue evenness are calculated from correlation matrices, whereas the other four metrics are calculated from covariance matrices. In all formulae \( \lambda \) are eigenvalues and \( n \) is the number of traits in the matrix (rank). *\( n_0 \) does not measure dimensionality per se, but eccentricity.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Cov/cor</th>
<th>Reference</th>
<th>Equation #</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>effective number of dimensions* ((n_b))</td>
<td>cov</td>
<td>[20]</td>
<td>#2 (pg 273)</td>
<td>( n_D = \sum_{i=1}^{n} \lambda_i / \lambda_i )</td>
</tr>
<tr>
<td>maximum evolvability ((e_{max}))</td>
<td>cov</td>
<td>[20]</td>
<td>#3 (pg 274)</td>
<td>( e_{max} = \sqrt{\lambda_i} )</td>
</tr>
<tr>
<td>total genetic variance ((v_T))</td>
<td>cov</td>
<td>[20]</td>
<td>#4 (pg 274)</td>
<td>( v_T = \sum_{i=1}^{n} \lambda_i )</td>
</tr>
<tr>
<td>average evolvability ((\bar{e}))</td>
<td>cov</td>
<td>[19]</td>
<td>#4 (pg 1206)</td>
<td>( \bar{e} = \sum \frac{\lambda_i}{n} )</td>
</tr>
<tr>
<td>eigenvalue variance ((\text{Var}(\lambda)))</td>
<td>cor</td>
<td>[49]</td>
<td>n/a (pg 158)</td>
<td>( \text{Var}(\lambda) = \frac{\sum_{i=1}^{n} (\lambda_i - 1)^2}{n} )</td>
</tr>
<tr>
<td>Relative eigenvalue variance ((\text{Var}_{rel}(\lambda)))</td>
<td>cor</td>
<td>[49]</td>
<td>n/a (pg 159)</td>
<td>( \text{Var}_{rel}(\lambda) = \frac{\text{Var}(\lambda)}{n - 1} )</td>
</tr>
</tbody>
</table>
| eigenvalue evenness \((E_i)\)               | cor     | [23]      | #3.2 (pg 1187)| \( E_i = -\sum_{i=1}^{n} \frac{\lambda_i \ln(\lambda_i)}{\ln(n)} \) \[
| where \( \tilde{\lambda}_i = |\lambda_i| / \sum_{j=1}^{n} |\lambda_j| \) |
Table 2. The main effects included in the final models for each analysis. (Effects of ‘trait type’ refer to life-history, morphology or sexual and ‘taxa’ to plant or animal. ‘Study type’ refers to field observation or experimental evolution. Random effects of ‘study’ and ‘species’ refer to models where an intercept was fitted to each species and study, and the random effect of ‘trait type:species’ indicates where both a species-level intercept and a species-level trait type effect were fitted.) Full sets of models can be found in the scripts and data on Dryad.

<table>
<thead>
<tr>
<th>measure</th>
<th>fixed effects</th>
<th>random effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rate (Animals)</td>
<td>trait type + study type</td>
<td>study + trait type:species</td>
</tr>
<tr>
<td>Rate (Plants)</td>
<td>trait type</td>
<td>species</td>
</tr>
<tr>
<td>$</td>
<td>\beta</td>
<td>\ $</td>
</tr>
<tr>
<td>(G) $nD$</td>
<td>trait type + taxon + trait no.</td>
<td>study</td>
</tr>
<tr>
<td>(G) $e_{max}$</td>
<td>trait type + taxon + trait no.</td>
<td>study + trait type:species</td>
</tr>
<tr>
<td>(G) $tgv$</td>
<td>trait type + taxon + trait no.</td>
<td>study + trait type:species</td>
</tr>
<tr>
<td>(G) $\bar{e}$</td>
<td>trait type + taxon + trait no.</td>
<td>study + trait type:species</td>
</tr>
<tr>
<td>(G) $\text{Var}(\lambda)$</td>
<td>trait type + taxon + trait no.</td>
<td>study</td>
</tr>
<tr>
<td>(G) $\text{Var}_{rel}(\lambda)$</td>
<td>trait type + taxon + trait no.</td>
<td>study</td>
</tr>
<tr>
<td>(G) $E_\lambda$</td>
<td>trait type + taxon + trait no.</td>
<td>study</td>
</tr>
</tbody>
</table>
Table 3. Summary statistics for estimates of the rate of evolutionary response, linear and quadratic selection gradients and measure capturing the size, shape and structure of G. (Statistics are reported by taxa and trait type, together with overall estimates across trait types and taxa. For each combination of taxa and trait type, the summary statistics for each measure are provided in the following order: posterior mean, posterior mode, lower and upper 95% credible intervals (in parenthesis) and sample size (in italics).)

<table>
<thead>
<tr>
<th>Measure</th>
<th>animals</th>
<th>overall</th>
<th>plants</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LH</td>
<td>M</td>
<td>SS</td>
</tr>
<tr>
<td>Rate (haldanes)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.00,0.19)</td>
<td>0.09</td>
<td>0.22</td>
<td>0.19</td>
</tr>
<tr>
<td>(0.02,0.22)</td>
<td>0.12</td>
<td>0.13</td>
<td>0.18</td>
</tr>
<tr>
<td></td>
<td>0.122</td>
<td>0.12</td>
<td>0.193</td>
</tr>
<tr>
<td></td>
<td>2.25</td>
<td>7</td>
<td>1667</td>
</tr>
<tr>
<td></td>
<td>781</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(G) nD</td>
<td>1.40</td>
<td>1.20</td>
<td>1.31</td>
</tr>
<tr>
<td>(0.76,1.5)</td>
<td>1.13</td>
<td>1.20</td>
<td>1.31</td>
</tr>
<tr>
<td></td>
<td>1.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(G) e&lt;sub&gt;max&lt;/sub&gt;</td>
<td>0.26</td>
<td>0.01</td>
<td>0.86</td>
</tr>
<tr>
<td>(0.1,0.4)</td>
<td>0.43</td>
<td>1.25</td>
<td>0.78</td>
</tr>
<tr>
<td></td>
<td>0.47</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(G) t&lt;sub&gt;gv&lt;/sub&gt;</td>
<td>8.62</td>
<td>25.41</td>
<td>14.32</td>
</tr>
<tr>
<td>(0.01,0.18)</td>
<td>17.16</td>
<td>23.57</td>
<td>21.19</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(G) ê</td>
<td>0.55</td>
<td>1.74</td>
<td>1.29</td>
</tr>
<tr>
<td>(0.2,0.94)</td>
<td>1.17</td>
<td>3.69</td>
<td>1.95</td>
</tr>
<tr>
<td></td>
<td>0.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>n (cov)</td>
<td>14</td>
<td>38</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(G) Var(λ&lt;sub&gt;n&lt;/sub&gt;)</td>
<td>0.39</td>
<td>0.79</td>
<td>1.19</td>
</tr>
<tr>
<td>(0.1,0.4)</td>
<td>0.36</td>
<td>0.43</td>
<td>0.50</td>
</tr>
<tr>
<td></td>
<td>0.45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(G) Var&lt;sub&gt;rel&lt;/sub&gt;(λ)</td>
<td>0.45</td>
<td>0.46</td>
<td>0.49</td>
</tr>
<tr>
<td>(0.15,0.56)</td>
<td>0.43</td>
<td>0.46</td>
<td>0.49</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(G) E&lt;sub&gt;λ&lt;sub&gt;2&lt;/sub&gt;&lt;/sub&gt;</td>
<td>0.76</td>
<td>0.70</td>
<td>0.67</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\text{n (cor)}$</td>
<td>0.78</td>
<td>0.77</td>
<td>0.69</td>
</tr>
<tr>
<td>----------------</td>
<td>------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td></td>
<td>(0.67,0.84)</td>
<td>(0.61,0.77)</td>
<td>(0.57,0.77)</td>
</tr>
<tr>
<td></td>
<td>42</td>
<td>82</td>
<td>27</td>
</tr>
</tbody>
</table>

762 Short Title for Page Headings: *Evolutionary rate as a function of selection & the G matrix*
Figure 2. Posterior means and 95% credible intervals for estimates of absolute rate of evolution (haldanes). Open points are for plants and filled points for animals. Trait types are life-history (LH), morphology (M) and sexually selected (S) and filled points are for animals and open points for plants (no data available for sexual traits in plants).
Figure 3. Posterior means and 95% credible intervals for estimates of standardized selection gradients (β) by trait type. Trait labels and taxa symbols are as in Figure 2.