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

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

\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 and Biochemistry, University of Bath, Claverton Down, Bath BA2 7AY, UK
\textsuperscript{3}College of Life and Environmental Sciences, Centre for Ecology and Conservation, University of Exeter, Tremough Campus, Penryn, Cornwall TR10 9EZ, UK

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 \((\Delta \mathbf{z} = \mathbf{G}\mathbf{b})\), which predicts evolutionary change for a suite of phenotypic traits \((\Delta \mathbf{z})\) as a product of directional selection acting on them \((\mathbf{b})\) and the genetic variance–covariance matrix for those traits \((\mathbf{G})\). Despite being empirically challenging to estimate, there are enough published estimates of \(\mathbf{G}\) and \(\mathbf{b}\) 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 some evidence that sexually selected traits exhibit faster rates of evolution compared with 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 \(\mathbf{G}\), 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.

1. Introduction

Predicting the rate and direction of phenotypic evolution remains a fundamental challenge in evolutionary biology [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- [9–11] and long-term [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 \(\Delta \mathbf{z} = \mathbf{G}\mathbf{b}\), where \(\Delta \mathbf{z}\) is the vector of responses in phenotypic means for the suite of traits, \(\mathbf{G}\) is the additive genetic variance–covariance matrix and \(\mathbf{b}\) 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 \(\mathbf{G}\) 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’ (LLER); directions in which the multivariate evolutionary response can proceed more rapidly than in others [15]. The presence of these LLERs can have a major influence in biasing the direction of evolutionary trajectories (figure 1; [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 [7,10,17,18,20,22,23,42].

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: $G$, $b$ and $\Delta z$ [5,17–19,21,24]. Unfortunately, few studies simultaneously estimate more than one of these components. The 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 benefited 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 references [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 with viability [1–4,33,38]. However, inferences from such studies are subject to methodological debate [5–8,35] and potentially publication biases [9–11,40]. In particular, there has been 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 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 with 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 & Sinchcombe [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 broader evolutionary conclusions [19,22,41].

We then examined how the 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 broader evolutionary conclusions [19,22,41].

2. Methods

All data and scripts containing our analyses can be downloaded either from DRYAD (doi:10.5061/dryad.g481c) or github (https://github.com/DworkinLab/Pitchers_PTRS2014).
(a) 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. palaeontologists), because 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 the use of 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 covary), 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 versus standard deviation) on estimates of selection [19,35], we have only included estimates standardized using the approach as advocated by Lande & Arnold [21], as this has been most 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: $V_A = h^2V_P$ and $cov_{n(a)} = r_{G_i}V_{A_{ij}}/V_{A_{ii}}$, where $V_A$ and $V_P$ are the additive genetic and phenotypic variances, $h^2$ is the narrow sense heritability and $r_{G_i}$ is the genetic correlation between traits $i$ and $j$. 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. Because 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 covariance 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 found to include cells with correlations more than one, and in these cases, we excluded the offending matrix.

(b) Defining trait categories and measures

Because 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 morphological 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 submatrices relating to traits within only 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 $B$ 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 used 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 measures we used explicitly accounted for the number of traits included in the matrix (i.e. $n_t$). While, in general, the number of traits seemed to have a small influence on these measures (figures 4 and 5), we also took several steps to account for these effects, such as including number of traits as a linear covariate in the models (below) and also examining the effects of scaling $n_t$ 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 $n_G$, 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’ ($\text{tgv}$), ‘maximum evolvability’ ($\text{emax}$) and ‘effective number of dimensions’ ($n_E$) and also Hansen & 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 ($E_\lambda$). Both sets of $G$ matrix measures are defined in table 1.

While we present the 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, because it is well known that different methods of scaling (i.e. mean-standardizing (co)variance matrices versus 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 the current standard
were fitted.) Full sets of models can be found in the scripts and data on Dryad.

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

animal. ‘Study type’ refers to field observation or experimental evolution. Random effects of ‘study’ and ‘species’ refer to models where an intercept was fitted

mates of standardized selection gradients (\(n_0\))

data with a meta-analysis, including random effects to account

meta-analyses, and we followed their lead in modelling selection

estimates of \(G\) also reported standard errors or confidence intervals

v. 2.15; [51]). A large proportion of studies reporting selection

ized linear-mixed-effect models using the MCMCglmm package

(c) Statistical analyses

expression of \(G\) used for response to selection [21], and rates calculated from correlation matrices would also not be directly comparable to those calculated from covariance matrices.

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 types: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>(\beta)</td>
<td>trait type + taxon + trait type (\times) taxon</td>
<td>study + species</td>
</tr>
<tr>
<td>(G) (n_0)</td>
<td>trait type + taxon + trait no.</td>
<td>study</td>
</tr>
<tr>
<td>(G) (e_{\text{max}})</td>
<td>trait type + taxon + trait no.</td>
<td>study + trait types:species</td>
</tr>
<tr>
<td>(G) (\text{tgv})</td>
<td>trait type + taxon + trait no.</td>
<td>study + trait types:species</td>
</tr>
<tr>
<td>(G) (\bar{e})</td>
<td>trait type + taxon + trait no.</td>
<td>study + trait types: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}_{\text{rel}}(\lambda))</td>
<td>trait type + taxon + trait no.</td>
<td>study</td>
</tr>
<tr>
<td>(G) (\bar{E}_\lambda)</td>
<td>trait type + taxon + trait no.</td>
<td>study</td>
</tr>
</tbody>
</table>

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 (DIC) values [52], and confirmed our selections by refitting the model set using reduced maximum-likelihood (lme4 package [53]) and comparing fits using Akaike and Bayesian information criterion scores (AIC/ BIC) and likelihood ratio tests 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 and github. Because we modelled the magnitude (absolute value) of our response variables, we used the folded normal distribution [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].

In total, we used 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

Table 1. \(G\)-matrix measures 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. \(n_0\) does not measure dimensionality per se, but eccentricity.

<table>
<thead>
<tr>
<th>measure</th>
<th>cov/cor</th>
<th>references</th>
<th>equation no.</th>
<th>formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>effective number of dimensions (^2) ((n_0))</td>
<td>cov</td>
<td>[20]</td>
<td>no. 2 (p. 273)</td>
<td>(n_0 = \sum_{i=1}^{n} \lambda_i / \lambda_1)</td>
</tr>
<tr>
<td>maximum evolvability ((\varepsilon_{\text{max}}))</td>
<td>cov</td>
<td>[20]</td>
<td>no. 3 (p. 274)</td>
<td>(\varepsilon_{\text{max}} = \sqrt{\lambda_1})</td>
</tr>
<tr>
<td>total genetic variance ((\nu_1))</td>
<td>cov</td>
<td>[20]</td>
<td>no. 4 (p. 274)</td>
<td>(\nu_1 = \sum_{i=1}^{n} \lambda_i)</td>
</tr>
<tr>
<td>average evolvability ((\bar{e}))</td>
<td>cov</td>
<td>[19]</td>
<td>no. 4 (p. 1206)</td>
<td>(\bar{e} = \sum_{i=1}^{n} \lambda_i / n)</td>
</tr>
<tr>
<td>eigenvalue variance ((\text{var}(\lambda)))</td>
<td>cor</td>
<td>[49]</td>
<td>n.a. (p. 158)</td>
<td>(\text{var}(\lambda) = \sum_{i=1}^{n} (\lambda_i - 1)^2 / n)</td>
</tr>
<tr>
<td>relative eigenvalue variance ((\text{var}_{\text{rel}}(\lambda)))</td>
<td>cor</td>
<td>[49]</td>
<td>n.a. (p. 159)</td>
<td>(\text{var}_{\text{rel}}(\lambda) = \text{var}(\lambda) / n - 1)</td>
</tr>
<tr>
<td>eigenvalue evenness ((E_\lambda))</td>
<td>cor</td>
<td>[23]</td>
<td>no. 3.2 (p. 1187)</td>
<td>(E_\lambda = -\sum_{i=1}^{n} \frac{\lambda_i \ln (\lambda_i)}{\ln (n)}), where (\hat{\lambda}_i =</td>
</tr>
</tbody>
</table>

\(^2\)In all formulae \(\lambda\) are eigenvalues and \(n\) is the number of traits in the matrix.
and animal rates separately. We had 776 estimates of $b$, 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.

### 3. Results

(a) 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 to 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 approximately 2.0 times as large (95% credible interval 0.7–4.8) as that for morphology, with only modest overlap of the 95% CIs for the two trait types (table 3). In animals, life history and morphology have similar estimates, but the posterior mean estimate for sexually selected traits is somewhat higher—1.5 times that for morphology (95% CI 0.5–6.9 times), and 1.5 times that for life history (95% CI 0.8–2.3 times). Furthermore, the 95% CIs for morphology do not include the estimate for sexually selected traits.

<table>
<thead>
<tr>
<th>measure</th>
<th>animals</th>
<th></th>
<th></th>
<th>plants</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LH</td>
<td>M</td>
<td>SS</td>
<td>overall</td>
<td>LH</td>
<td>M</td>
</tr>
<tr>
<td>rate (haldanes)</td>
<td>0.12</td>
<td>0.13</td>
<td>0.18</td>
<td>0.13</td>
<td>0.3</td>
<td>0.15</td>
</tr>
<tr>
<td>(0.02,0.22)</td>
<td>(0.09,0.17)</td>
<td>(0.10,0.26)</td>
<td>(0.08,0.17)</td>
<td>(0.18,0.42)</td>
<td>(0.05,0.25)</td>
<td>—</td>
</tr>
<tr>
<td>$</td>
<td>\beta</td>
<td></td>
<td>0.09</td>
<td>0.22</td>
<td>0.19</td>
<td>0.21</td>
</tr>
<tr>
<td>(0.00,0.19)</td>
<td>(0.16,0.28)</td>
<td>(0.11,0.27)</td>
<td>(0.17,0.26)</td>
<td>(0.22,0.41)</td>
<td>(0.09,0.36)</td>
<td>(0.00,0.37)</td>
</tr>
<tr>
<td>$G_n$</td>
<td>1.13</td>
<td>1.20</td>
<td>1.31</td>
<td>1.53</td>
<td>1.23</td>
<td>1.29</td>
</tr>
<tr>
<td>(0.76,1.5)</td>
<td>(0.82,1.54)</td>
<td>(0.85,1.82)</td>
<td>(1.39,1.67)</td>
<td>(0.77,1.65)</td>
<td>(0.86,1.71)</td>
<td>(0.99,1.81)</td>
</tr>
<tr>
<td>$G_{e_{\max}}$</td>
<td>0.43</td>
<td>1.25</td>
<td>0.78</td>
<td>0.61</td>
<td>0.59</td>
<td>0.91</td>
</tr>
<tr>
<td>(0.10,0.4)</td>
<td>(0.8,1.83)</td>
<td>(0.8,1.88)</td>
<td>(0.26,0.97)</td>
<td>(0.13,0.97)</td>
<td>(0.22,0.22)</td>
<td>(0.12,0.28)</td>
</tr>
<tr>
<td>(1.61,18.11)</td>
<td>(0.3,52.47)</td>
<td>(0.3,108.9)</td>
<td>(0.7,7.57)</td>
<td>(0.01,19.64)</td>
<td>(0.01,50.48)</td>
<td>(0.01,22.86)</td>
</tr>
<tr>
<td>$G_{\bar{e}}$</td>
<td>1.17</td>
<td>3.69</td>
<td>1.95</td>
<td>0.48</td>
<td>1.29</td>
<td>3.40</td>
</tr>
<tr>
<td>(0.2,0.94)</td>
<td>(0.01,8.55)</td>
<td>(0.4,7.57)</td>
<td>(0.15,0.15)</td>
<td>(0.3,0.3)</td>
<td>(0.7,0.73)</td>
<td>(0.3,0.62)</td>
</tr>
<tr>
<td>$n$ (cov)</td>
<td>14</td>
<td>38</td>
<td>10</td>
<td>81</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>$G_{\text{var}(\lambda)}$</td>
<td>0.43</td>
<td>0.60</td>
<td>0.80</td>
<td>1.40</td>
<td>0.63</td>
<td>0.49</td>
</tr>
<tr>
<td>(0.10,0.4)</td>
<td>(0.13,0.68)</td>
<td>(0.9,2.16)</td>
<td>(0.14,0.2)</td>
<td>(0.1,0.17)</td>
<td>(0.1,0.14)</td>
<td></td>
</tr>
<tr>
<td>$G_{\text{var}(\lambda)_{rel}}$</td>
<td>0.36</td>
<td>0.43</td>
<td>0.50</td>
<td>0.32</td>
<td>0.16</td>
<td>0.23</td>
</tr>
<tr>
<td>(0.15,0.56)</td>
<td>(0.22,0.63)</td>
<td>(0.25,0.75)</td>
<td>(0.23,0.41)</td>
<td>(0.03,0.33)</td>
<td>(0.04,0.42)</td>
<td>(0.06,0.53)</td>
</tr>
<tr>
<td>$G_{\bar{\lambda}}$</td>
<td>0.76</td>
<td>0.70</td>
<td>0.67</td>
<td>0.73</td>
<td>0.86</td>
<td>0.79</td>
</tr>
<tr>
<td>(0.67,0.84)</td>
<td>(0.61,0.77)</td>
<td>(0.57,0.77)</td>
<td>(0.70,0.77)</td>
<td>(0.77,0.94)</td>
<td>(0.71,0.87)</td>
<td>(0.68,0.86)</td>
</tr>
<tr>
<td>$n$ (cor)</td>
<td>42</td>
<td>82</td>
<td>27</td>
<td>221</td>
<td>14</td>
<td>26</td>
</tr>
</tbody>
</table>

Table 3. Summary statistics for estimates of the rate of evolutionary response, linear and quadratic selection gradients and measures 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 parentheses) and sample size (in italics). LH, life history; M, morphology; SS, sexually selected.
traits, though those for life history do. Despite this, model support from various measures (AIC, BIC and DIC) is inconsistent about the overall support of trait types for the animal data improving model fit. Overall, these results suggest similar rates of evolution for morphology in both plants and animals, with higher rates for life-history traits in plants and possibly for sexually selected traits in animals.

(b) Standardized 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 laboratory 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 = 0.31–0.56) that for morphology, and 0.49 times (95% CI = 0.47–0.52) that for sexually selected traits (figure 3).

(c) 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 (table 1), we report \( t_{G^D}, e_{\text{max}} \) and \( \bar{r} \) can be thought of as measures of magnitude, whereas \( n_{2D}, \var(\lambda), \var_{\text{rel}}(\lambda) \) and \( E_{\lambda} \) 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 and figures 4 and 5), because \( t_{G^D}, e_{\text{max}} \) and \( \bar{r} \) are all intercorrelated (\( r > 0.96 \) 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 in retrospect 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 \( \var(\lambda) \) and \( \var_{\text{rel}}(\lambda) \) are strongly correlated with each other (\( r = 0.87 \)), and negatively correlated with \( E_{\lambda} \) (\( r = -0.32 \) and \( -0.55 \), respectively). Although \( E_{\lambda} \) 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] \( n_{2D} \) (\( r = 0.82 \)).

(d) 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 \( e_{\text{max}}, t_{G^D} \) and \( \bar{r} \). 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 \( n_{2D} \) 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 6d), but once
again, there is wide overlap among credible intervals, indicating low confidence in this trend. While this is for the inclusion of trait number as a linear covariate, similar results were obtained when $n_D$ was scaled directly by trait number (electronic supplementary material, figure S1).

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 7a). The opposite trend was present in estimates for $\text{var}(\lambda)$ with the estimates for animals being somewhat lower than those for plants (figure 7b). The wide overlap of credible intervals indicates low confidence in both these trends however. Finally, our estimates for $E_{\lambda}$ show a decreasing trend from life-history to morphology to sexually selected traits in both plants and animals, again with higher estimates for plants than for animals (figure 7c).

4. Discussion

Predicting the rate and direction of phenotypic evolution is a fundamental challenge in evolutionary genetics [1–4,54], and the multivariate breeders’ equation is a key tool. Estimates of $G$, selection and of response are available in the literature from many systems (though rarely reported together). Here, we have integrated these data to ask whether some traits evolve more rapidly than others, and whether differences associate with 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 some evidence that in animals—though not plants—sexual traits evolve faster than morphological traits. We find no evidence that this is 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 $G$.

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’ ($e_{\text{max}}$) and ‘effective number of dimensions’ ($n_D$) [20], the first eigenvalue of $G$ ($g_{\text{max}}$), ‘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 ($n$). Figures in the lower off-diagonal are pairwise correlations between the measures.
among trait types, though this fails to provide an explanation for any increased rates of evolution.

(a) Similarities among measures of the size and structure of \( G \)

We examined a number of the measures that have been proposed to assess the size, shape and structure of \( G \) (table 1). Many of these measures have considerable shared information (figures 4 and 5). Broadly, one group expresses the magnitude of \( G \) and a second relates to the evenness/variance of the eigenvalues, or eccentricity of \( G \). While there may be particular instances where these measures result in widely divergent estimates, with respect to the empirical estimates we have collated, the marginal benefits of using all of them are an illustration of diminishing returns. It remains possible that subtle differences among these measures may provide important insights into the structure of \( G \) in the future. We speculate that one potential use (which would require considerable additional research) may be analogous to the population geneticists’ use of the parameter Tajima’s \( D \), which is a scaled measure of two different estimates of the population mutation rate, \( 4N_e\mu \).

One surprising observation that emerges from our results is that the number of traits \( n \) used to estimate \( G \) is not well correlated with any of the measures we used. One potential explanation for this is that the magnitude 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 \( G \)), consistent with previous studies [9,20,23]. It is known that estimating \( G \) can be difficult and insufficient sampling at the level of families can inflate the magnitude of the principal eigenvalue, at the expense of the minor eigenvalues [55,56]. However, we saw no signal of such an effect from this database with any measures that capture eccentricity for \( G \) (electronic supplementary material, figures S4 and S5). As we did not have the raw data to recompute \( G \) in a consistent framework, it is unclear how substantial this bias might be.

It is well known that scaling trait values by the mean versus the standard deviation can have profound impacts.

Figure 5. Pairs plot to illustrate the relationships between measures used to describe the structure of \( G \) expressed as correlation matrices. Measures are ‘relative eigenvalue variance’ \( \text{var}_{rel}(\lambda) \) [49], ‘eigenvalue evenness’ \( E_\lambda \) [23], ‘eigenvalue variance’ \( \text{var}(\lambda) \) [49], the first eigenvalue of \( G \) \( (g_{\text{max}}) \) and the number of traits included in the matrix \( n \). 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 characterize G matrices expressed as covariances (see methods section): (a) ‘maximum evolvability’ ($e_{\text{max}}$), (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); filled points are for animals and open points for plants.

Figure 7. Posterior means and 95% credible intervals for the four measures used to characterize G matrices expressed as correlations (see Methods section): (a) ‘relative eigenvalue variance’ ($\text{var}_{rel}(\lambda)$), (b) ‘eigenvalue variance’ ($\text{var}(\lambda)$) and (c) ‘eigenvalue evenness’ ($E_\lambda$). Trait labels and taxon symbols are as in figure 6.
on univariate measures such as heritability. Likewise, this would be expected for multivariate extensions such as $G$ and measures extracted from them as used here. Unfortunately, in many instances, the vector of trait means was unavailable, and thus our analysis for mean-scaled $G$ is a subset of that for the correlation matrices.

(b) Rates of evolution vary among traits

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 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 larger study, Kinnison & Hendry [31] showed that the frequency distribution of evolutionary rates measured in haldanes is lognormal (i.e. many slow rates and few fast rates, median haldanes $= 5.8 \times 10^{-5}$) 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 lognormal and that the median rate across trait types and taxa was similar (median haldanes $= 7.6 \times 10^{-5}$) to that reported in Kinninson & Hendry [31]. We found little evidence to suggest that the evolutionary rates of life-history and morphological traits differed in animals, though there is evidence for faster rates in plant life history. Our findings provide some 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]. It is worth noting that we used a different method for scaling data, as well as the inclusion of laboratory-based studies of evolutionary rates, which differs from some other recent studies such as Uyeda et al. [46]. Future work examining how different methods of examining rate and the inclusion of laboratory versus field samples influence the overall observed pattern is warranted.

(c) The strength of selection varies among traits

Reviews synthesizing estimates of selection are extensive [33–39]. In their seminal review, Kingsolver et al. [33] found that the frequency distributions of linear and quadratic selection gradients were exponential and generally symmetrical around zero. This suggests that stabilizing and disruptive selection occur with equal frequency and with similar strength in nature. Kingsolver et al. [33] also found that the magnitude of linear selection was on average greater for morphological rather than life-history traits. The most recent review [38] containing an updated dataset 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 synthesis [38], we found that the distribution of linear and quadratic selection gradients were exponential. 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)). There has been much discussion on the general limitations of using selection gradients in synthetic reviews [33,35,37,38], and these arguments undoubtedly also apply to our study. However, as most of these limitations are inherent to both studies, they are unlikely to explain the observed differences. Furthermore, we used the same Bayesian framework as Kingsolver et al. [38], so it is unlikely that our analytical approach 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. By contrast, we distinguished only 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 behavioral traits. Thus, there are likely to be some differences in how selection gradients are distributed among categories in our analyses compared with those in Kingsolver et al. [38]. Irrespective of the underlying reasons for these differences, we find little evidence for differences in the magnitude of selection gradients across trait types and taxa.

(d) Evolutionary response and the structure of $G$

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 directions—known as ‘LLER’ [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. 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 intercorrelated 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.

Our 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 constraint 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 [48] and empirical [71] evidence suggesting that floral morphology is indeed subjected to sexual selection. Unfortunately though, there are currently no available 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.
(e) The effect of trait scaling

Researchers need to remain mindful that decisions about measurement scaling 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]. In this paper, we have attempted to present a clear picture of the patterns present in the currently available data, but it is important to acknowledge the known shortcomings of that data. This is not to understate the difficulty of maintaining comparability among studies wherein the appropriate scales might be different [6,35,72]. To illustrate the problem, how best to compare morphological data comprising linear measurements with life-history data where there may be no natural zero value? As a field, our inferences about selection and the response to selection will be more meaningful the more clearly we can address these issues.

5. Conclusion

Collectively, our results suggest that the higher rate of evolution observed for sexual traits in animals is only weakly associated with the scalar measures summarizing G for these traits, and we do not find stronger selection. However, as our dataset is based on derived estimates there are a number of inevitable limitations that apply to our findings. First, there are limitations with using the matrix structure measures (\(n_{xy}\), \(E_s\), \(\text{var}(A)\) or \(\text{var}_{e}(A)\)) to capture the dimensionality of \(G\) [20]. Although these measures 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 [61,63] have been taken to directly estimate the dimensionality of \(G\) [63], though such studies have found both populations that have evolutionary access to all dimensions of \(G\) [63] and others that are constrained by LLERs [61,74]. Second, our analysis does not consider the alignment between the vectors of selection and G. LLERs only constrain the response to selection when they are poorly aligned with vectors of selection [26,28,64]. These limitations can be resolved only by further analysis of the raw datasets from 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 [56,61,63,75–77]. Future studies would greatly benefit from researchers publishing raw datasets in open repositories [78] and we encourage researchers to do so.

Acknowledgements. We thank S. F. McDaniel, R. L. Rodriguez and W. U. Blankenhorn for contributing genetic data and to R. Snook, S. Chenoweth, T. Chapman, R. Firman and M. Reuter for data on experimental evolution. I. Stott, F. C. Ingleby and three anonymous reviewers provided comments on earlier drafts of this manuscript, and A. Haber proofread the associated scripts.

Data accessibility. Our database (with all associated analyses) can be found at DRYAD (doi:10.5061/dryad.g68tsc), or on github (https://github.com/DworkinLab/Pitchers_PTRS2014). Funding statement. J.H., T.T. and J.B.W. were supported by NERC, J.H. and J.B.W. by the BBfSRC and J.H. by a University Royal Society Fellowship. W.P. was supported by an NERC studentship (awarded to J.H. and T.T.) and I.D. and W.P. were funded by NIH grant no. 1R01GM094424.

References

23. Agrawal AF, Stinchcombe JR. 2009 How much do genetic covariances alter the rate of adaptation?


