Alternative mutational architectures can produce identical $\mathbf{M}$-matrices but lead to different patterns of macroevolutionary divergence

Daohan Jiang[1] & Matt Pennell[1,2,*]

[1]*Department of Quantitative and Computational Biology, University of Southern California, USA*
[2]*Department of Biological Sciences, University of Southern California, USA*
[*]Corresponding author: mpennell@usc.edu

Abstract

Explaining macroevolutionary divergence in light of population genetics requires understanding the extent to which the patterns of mutational input contribute to long-term trends. In the context of quantitative traits, mutational input is typically described by the mutational variance-covariance matrix, or the $\mathbf{M}$-matrix, which summarizes phenotypic variances and covariances introduced by new mutations per generation. However, as a summary statistic, the $\mathbf{M}$-matrix does not fully capture all the relevant information from the underlying mutational architecture, and there exist infinitely many possible underlying mutational architectures that give rise to the same $\mathbf{M}$-matrix. Using simulations under a strong-selection-weak-mutation regime, we demonstrate alternative mutational architectures underlying the same $\mathbf{M}$-matrix can lead to different levels of constraint on evolution, and different patterns of alignment between $\mathbf{M}$ and the evolutionary variance-covariance matrix, or the $\mathbf{R}$-matrix. Our results reveal that aspects of mutational input not reflected by the $\mathbf{M}$-matrix can have a profound impact on long-term evolution, and suggest it is important to take them into account in order to connect macroevolutionary patterns to microevolutionary mechanisms.

Keywords: $\mathbf{M}$-matrix, macroevolutionary divergence, phenotypic evolution, pleiotropy

Introduction

A fundamental problem in evolutionary biology is explaining phenotypic diversity across species in light of population genetics [Arnold et al., 2001, Hansen and Houle, 2004, Harmon et al., 2021, Rolland et al., 2023] At least in some taxa, macroevolutionary divergence, captured by the variances and covariance in phenotypes among lineages on a phylogenetic tree, reflects the covariance structure of the additive genetic variance within lineages [Ackermann and Cheverud, 2004, Bolstad et al., 2014, McGlothlin et al., 2018, Opedal et al., 2023, Porto et al., 2015, Schluter, 1996]. While we expect the short term response to selection to be in the direction with the most additive genetic variance in the phenotypic space [Arnold, 1992, Lande, 1976, 1979, Walsh and Blows, 2009], it is intriguing that this appears to extend to longer term trends. Selection will tend to erode the additive genetic variance relatively quickly (at least relative to macroevolutionary timescales) so the structure of additive genetic variance must be maintained by some combination of drift, correlated selection, and correlated effects of mutations [Arnold, 1992, Arnold et al., 2008, Chakraborty and Nei, 1982, Cheverud, 1984, Jones et al., 2003, 2007, Lande, 1979, 1980, McGuigan, 2006, Svensson, 2022, Svensson et al., 2021]. For instance, strong alignment between the patterns genetic variation and macroevolutionary divergence either can be explained either by the absence of selection, or by a model where both are shaped by selection in the same way (i.e., traits or trait combinations subject to stronger selection have both lower genetic variances and lower evolutionary rates). The former explanation requires that the pattern of genetic variance reflects the patterns of mutational input — and this will be the focus of this paper.

That structure of mutational input is constraining the course of long-term phenotypic evolution is a long-standing and controversial hypothesis [Gould, 1980, Nei, 2013, Stoltzfus, 2021] but is notoriously difficult to test as the necessary data are very challenging to collect. A pioneering study [Houle et al., 2017] sought to take this on and test whether the observed correlations between the pattern of within-species additive genetic (co)variance (as encapsulated by the $\mathbf{G}$-matrix, hereafter $\mathbf{G}$) and the evolutionary phenotypic (co)variance (as encapsulated by the $\mathbf{R}$-matrix, hereafter $\mathbf{R}$) could be explained by mutations. Using wings of drosophilid flies as a model system, Houle et al. [2017] conducted laborious mutation accumulation (MA) experiments [reviewed in Halligan and Keightley, 2009] to estimate the mutational variance-covariance matrix (the $\mathbf{M}$-matrix, hereafter $\mathbf{M}$). The $\mathbf{M}$-matrix summarizes the amount and correlational structure of mutational input: each diagonal element of $\mathbf{M}$ represents a trait's mutational variance (i.e., phenotypic vari-

ance introduced by new mutations per unit time, usually presented on a per-generation basis), and each off-diagonal element represents the mutational covariance (i.e., phenotypic covariance introduced by new mutations per generation) between two traits. They used genetic crosses to estimate $\mathbf{G}$ and collected phylogenetic data to estimate $\mathbf{R}$. They then tested whether $\mathbf{M}$, $\mathbf{G}$, and $\mathbf{R}$ were aligned — the level of similarity between the matrices can be evaluated by comparing variances in two matrices on a pre-specified set of orthogonal dimensions in the phenotypic space quantified by the correlation between the variances. Again, a positive result would be consistent with the patterns of mutational input driving long-term evolution, and this alignment is exactly what Houle et al. [2017] reported. However, the strength of the scaling relationship was likely mis-estimated [Jiang and Zhang, 2020] and it remains an open problem as to what the best estimate of the scaling relationship is and how best to interpret the estimate [Houle et al., 2020, Jiang and Zhang, 2020, Opedal et al., 2023]. Though beyond the scope of this study, it is worth noting that such analyses of alignment could be confounded if $\mathbf{M}$ itself has evolved through time as a result of second-order selection [Jones et al., 2007, 2014]. As it is often not possible evaluate the role $\mathbf{M}$'s evolution on the alignment because $\mathbf{M}$ estimates from most taxa are unavailable, caution is necessary when interpreting the alignment.

Our purpose in this paper is not to delve into the analysis of this particular dataset; rather we will argue that the core predictions and interpretations of these tests — and future tests like them — are not as straightforward as they appear. The conceptual difficulty of testing for mutation-driven macroevolution lies in the fact that $\mathbf{M}$ is an insufficient summary statistic for the mutational architecture. By definition, the mutational variance is a product of the mutation rate and variance of mutations' effect on the trait; similarly, the mutational covariance of a given pair of traits is a product of the rate of pleiotropic mutations affecting both traits and the covariance of the mutations' effects on two traits [Hansen, 2006, Lynch and Hill, 1986]. That is, the same $\mathbf{M}$ can result from different combinations of parameters; importantly, these different combinations can have different dispositional effects on evolution. For example, the covariance is zero either when there are no pleiotropic mutations, or when some or all mutations are pleiotropic but pleiotropic mutations' effects on different traits are uncorrelated. As pleiotropic mutations tend to be more deleterious, phenotypic evolution is more constrained in the latter case, leading to lower evolutionary rate [Chevin et al., 2010, McGuigan et al., 2014, Orr, 2000, Wagner and Zhang, 2011], though this may not be true under all conditions as an intermediate level of pleiotropy could facilitate adaptation under certain conditions [Rennison and Peichel, 2022, Wang et al., 2010].

Consequently, in order to make sense of the connection between $\mathbf{M}$ and macroevolutionary patterns, strong assumptions about the mutational architectures have to be made. In particular, models for the evolution of multivariate traits often assume universal pleiotropy, where every mutation affects every trait. For instance, Fisher's Geometric Model makes this assumption, which leads to the prediction that mutations with smaller effect sizes are more likely to be adaptive and that there is a "cost of complexity" as adaptation is slower when there is a greater number of phenotypic dimensions [Fisher, 1930, Orr, 2000, Tenaillon, 2014, Welch and Waxman, 2003]. Similarly, in a series of modeling studies, Jones et al. [2007] and Jones et al. [2014] assumed universal pleiotropy when modeling the evolution of the mutational architecture under second-order selection, making the effect size correlation being the only evolvable aspect of mutational architecture; it is unknown whether the mutational architecture would evolve differently if the assumption of universal pleiotropy is relaxed. The degree to which such assumptions are reasonable remains an open question in quantitative genetics [Boyle et al., 2017, Hill and Zhang, 2012a,b, Paaby and Rockman, 2013, Wagner and Zhang, 2011, Zhang, 2023]. Some studies have found that each gene or mutation typically affects only a small subset of traits [Ho and Zhang, 2014, Wagner et al., 2008, Wang et al., 2010]. Others argue that pleiotropy is more pervasive and that many empirical studies underestimate the prevalence of pleiotropy due to technical issues [Hill and Zhang, 2012b]. Proponents of the recently proposed "omnigenic" model argue the weight of evidence from genetic mapping studies (particularly in the context of human diseases) indicate that each mutation can affect a large number of traits because of properties of the regulatory network [Boyle et al., 2017, Liu et al., 2019]. One way to make sense of these mixed results is to conclude that it is likely each individual gene or mutation usually has major effect on only a small set of traits, but may have small effects on many more traits — in other words, pleiotropy is widespread but the distribution of effect sizes is highly skewed.

To demonstrate how a myriad of different mutational architectures can produce the same $\mathbf{M}$, let us consider two quantitative traits z_1 and z_2. We note that our primary interest in this paper is the alignment between $\mathbf{M}$ and $\mathbf{G}$ or $\mathbf{R}$ and this is most interesting to study when the number of phenotypic dimensions is larger than two; however, the 2 trait case is useful for demonstrating the general principle. Assuming only additive effects, there are three types of mutations: mutations that affect z_1 only, mutations that affect z_2 only, and pleiotropic mutations that affect both traits. Let us assume that the effect size of mutations that exclusively influence a single trait are drawn from normal distributions, $\mathcal{N}(0, \sigma_{e.1})$, and $\mathcal{N}(0, \sigma_{e.2})$,

respectively, and the effect size of pleiotropic mutations follows a bivariate normal distribution characterized by a covariance matrix:

$$\mathbf{M}_P = \begin{bmatrix} \sigma_{p.1}^2 & \rho\sigma_{p.1}\sigma_{p.2} \\ \rho\sigma_{p.1}\sigma_{p.2} & \sigma_{p.2}^2 \end{bmatrix}. \tag{1}$$

Here, $\sigma_{p.1}$ and $\sigma_{p.2}$ are the standard deviations of pleiotropic mutations' effects on z_1 and z_2, respectively, and ρ denotes the correlation between the mutations' effects on z_1 and their effects on z_2. The $\mathbf{M}$ matrix is obtained by combining the (co)variances resulting from pleiotropic and non-pleiotropic mutations:

$$\mathbf{M} = U_1 \begin{bmatrix} \sigma_{e.1}^2 & 0 \\ 0 & 0 \end{bmatrix} + U_2 \begin{bmatrix} 0 & 0 \\ 0 & \sigma_{e.2}^2 \end{bmatrix} + U_P \begin{bmatrix} \sigma_{p.1}^2 & \rho\sigma_{p.1}\sigma_{p.2} \\ \rho\sigma_{p.1}\sigma_{p.2} & \sigma_{p.2}^2 \end{bmatrix}$$

Here, U_1 represents the rate of mutations that affect z_1 only, U_2 represents the rate of mutations that affect z_2 only, and U_P represents the rate of pleiotropic mutations. Note that the three mutation types can be perceived as corresponding to three classes of loci, and each mutation rate parameter (U_1, U_2, or U_P) can be interpreted as a product of the number of loci in the corresponding class and the corresponding class-specific per-locus mutation rate.

Consequently, the first term in the above equation represents mutational variance resulting from mutations that affect z_1 only, the second term represents mutational variance resulting from mutations that affect z_2 only, and the last term represents mutational (co)variances resulting from pleiotropic mutations. Adding up the tree terms in the previous equation gives:

$$\mathbf{M} = \begin{bmatrix} U_1\sigma_{e.1}^2 + U_P\sigma_{p.1}^2 & U_P\rho\sigma_{p.1}\sigma_{p.2} \\ U_P\rho\sigma_{p.1}\sigma_{p.2} & U_2\sigma_{e.2}^2 + U_P\sigma_{p.2}^2 \end{bmatrix} \tag{2}$$

We denote to a particular vector of values $\mathbb{P}$, where $\mathbb{P} : \{U_1, \sigma_{e.1}, U_2, \sigma_{e.2}, U_P, \rho, \sigma_{p.1}, \sigma_{p.2}\}$ hereafter for convenience. This model can be easily extended to phenotypic spaces with a higher number of dimensions; for an n-dimensional phenotypic space, the number of possible mutation types is equal to $\sum_{i=1}^{n} \binom{n}{i}$.

As shown in equation (2), each entry of $\mathbf{M}$ is contributed by multiple quantities. Thus, it is possible that different $\mathbb{P}$ give rise to the same $\mathbf{M}$ matrix. To see how we can manipulate the parameters while holding $\mathbf{M}$ constant, let U_P, ρ, $\sigma_{p.1}$, and $\sigma_{p.2}$ each be multiplied by a rescaling coefficient, such that they become $C_P U_P$, $C_\rho \rho$, $C_{p.1}\sigma_{p.1}$, and $C_{p.2}\sigma_{p.2}$, respectively, where $C_P C_\rho C_{p.1} C_{p.2} = 1$ and $C_\rho < 1/|\rho|$. Let us multiply $U_1\sigma_{e.1}^2$

and $U_2\sigma_{e.2}^2$ by rescaling coefficients C_1 and C_2, respectively, to keep the mutational variances unchanged:

$$\begin{cases} \mathbf{M}[1,1] = C_1U_1\sigma_{e.1}^2 + C_PC_{p.1}^2U_P\sigma_{p.1}^2 & = U_1\sigma_{e.1}^2 + U_P\sigma_{p.1}^2 \\ \mathbf{M}[2,2] = C_2U_2\sigma_{e.1}^2 + C_PC_{p.2}^2U_P\sigma_{p.2}^2 & = U_2\sigma_{e.2}^2 + U_P\sigma_{p.2}^2 \end{cases}$$

Solving the above equations gives

$$\begin{cases} C_1 = \dfrac{U_1\sigma_{e.1}^2 + (1 - C_PC_{p.1}^2)U_P\sigma_{p.1}^2}{U_1\sigma_{e.1}^2} \\ C_2 = \dfrac{U_2\sigma_{e.2}^2 + (1 - C_PC_{p.2}^2)U_P\sigma_{p.2}^2}{U_2\sigma_{e.2}^2} \end{cases} \tag{3}$$

C_1 and C_2 must be non-negative, as no mutation rate or standard deviation can be negative. Therefore, C_1 and C_2 can only be solved if

$$\begin{cases} U_1\sigma_{e.1}^2 + (1 - C_PC_{p.1}^2)U_P\sigma_{p.1}^2 > 0 \\ U_2\sigma_{e.2}^2 + (1 - C_PC_{p.2}^2)U_P\sigma_{p.2}^2 > 0 \end{cases}$$

Solving the above system of inequities gives

$$\begin{cases} C_PC_{p.1}^2 < \dfrac{U_1\sigma_{e.1}^2}{U_P\sigma_{p.1}^2} + 1 \\ C_PC_{p.2}b^2 < \dfrac{U_2\sigma_{e.2}^2}{U_P\sigma_{p.2}^2} + 1 \end{cases} \tag{4}$$

Hence, given $\mathbf{M}$, certain combinations of C_P, C_ρ, $C_{p.1}$, and $C_{p.2}$ are guaranteed to alter the mutational variances. Biologically, if the portion of mutational variance attributable to pleiotropic mutations gets too high, it would be impossible to keep the total mutational variance unchanged by reducing the portion contributed by non-pleiotropic mutations.

Given that C_1 can be solved, the change to $U_1\sigma_{e.1}^2$ can be done by altering U_1, $\sigma_{e.1}$, or both. Thus, for any given combination of C_P, C_ρ, $C_{p.1}$, and $C_{p.2}$, there exists infinitely many ways to adjust $U_1\sigma_{e.1}^2$ to keep $\mathbf{M}$ unchanged. Similarly, there are also infinitely many ways to adjust $U_2\sigma_{e.2}^2$. Hence, there exists infinitely many unique $\mathbb{P}$ that give rise to the same $\mathbf{M}$.

As noted above, we will leave the two phenotypic case for our simulations and consider a phenotypic space with higher dimensions to illustrate that the alignment between $\mathbf{M}$ and $\mathbf{R}$ depends not just on $\mathbf{M}$ but

on the underlying mutational architecture. Since there is a many-to-one mapping between $\mathbb{P}$ and $\mathbf{M}$, one cannot make clear predictions as to the expected relationship between $\mathbf{M}$, $\mathbf{G}$, and $\mathbf{R}$ or properly interpret their alignment without additional information or assumptions.

Materials and Methods

Mutational architectures

We considered a 10-dimensional phenotypic space and conducted simulations of evolution of the ten traits. The $\mathbf{M}$-matrix considered here, $\mathbf{M}_{10}$, is a 10×10 diagonal matrix with diagonal elements (i.e., traits 1-10, z_1,...,z_{10} hereafter) equal to $\exp(-9)$, $\exp(-8)$, $\exp(-7)$, ..., $\exp(0)$ and all off-diagonal elements equal to zero. Although the assumption of zero mutational covariances appears restrictive, a matrix with zero off-diagonal elements can represent the orthogonalized form of an arbitrary $\mathbf{M}$-matrix (i.e., $\mathbf{K}^{\mathrm{T}}\mathbf{M}\mathbf{K}$, $\mathbf{K}$ being the eigenvector matrix of the "original" $\mathbf{M}$), in which case z_1,...,z_{10} represent linear combinations of the "original" traits.

Since the number of mutation parameters is large (1024 in total), we focused on several special cases of interest, where the per-trait mutation rate (i.e., rate of mutations that affect each individual trait) and the distribution of mutational effect on each individual trait remained the same but the total mutation rate and prevalence of pleiotropy varied. We first considered a mutational architecture characterized by a lack of pleiotropy (i.e., every mutation affects one trait only), and one characterized by universal pleiotropy (i.e., every mutation affects all traits). In the first case, there are ten types of mutations (mutations affecting z_1, ..., mutations affecting z_{10}) and 20 underlying parameters (U_1, ..., U_{10}, $\sigma_{e.1}$, ..., σ_{10}). In our simulations, we had $U_1 = \ldots = U_{10} = 1$ (i.e., total mutation rate $U = U_1 + \cdots + U_{10} = 10$) and $\sigma^2_{e.1} = \mathbf{M}_{10}[1,1]$, ..., $\sigma^2_{10} = \mathbf{M}_{10}[10,10]$. In the case of universal pleiotropy, there is only one type of mutation, of which the rate was set to be 1 (i.e., total mutation rate $U = 1$), and the effect of mutation on traits follows a multivariate normal distribution corresponding to $\mathbf{M}_{10}$. The overall mutation rate in this scenario is 10% of that in the first one, but the per-trait mutation rate is the same.

As modularity is a property of the mutational architecture that is thought to have a profound impact on the evolution of multivariate traits [De Visser et al., 2003, Goswami et al., 2014, Melo and Marroig,

2015, Wagner, 1996, Wagner and Altenberg, 1996, Wagner et al., 2007, Wagner and Zhang, 2011], we also considered two mutational architectures with modular structures. We divided the ten mentioned above traits into two modules, each containing five traits. We assumed that a mutation can affect only one module, and that a mutation affecting a given module affects all traits in this module with effects on different traits uncorrelated. Specifically, we considered two modular organizations: in the first mutational architecture, the five traits with the smallest $\mathbf{M}$ variances (i.e., $z_1, \ldots, z_5$) are in the same module while the rest (i.e., $z_6, \ldots, z_10$) are in the other module; in the second case, z_1, z_3, z_5, z_7, and z_9 are in one module while the rest belong to the other such that the two modules have relatively similar total $\mathbf{M}$ variances. In both cases, there are two classes of mutations, each having a mutation rate of 1; thus, the total mutation rate is twice of that in the case of universal pleiotropy, and 20% of that in the case with no pleiotropy.

Mapping between phenotype and fitness

We considered a multivariate Gaussian fitness function, which is described by a covariance matrix $\mathbf{S}$. Each diagonal element of $\mathbf{S}$ is the width of an individual trait's fitness function (i.e., variance of a normal distribution), and off-diagonal elements represent correlational selection for relationships between traits [Arnold et al., 2008, 2001].

To calculate fitness given the n-dimensional phenotype $\vec{z}$, we first calculate its distance to the optimal phenotype $\vec{o}$:

$$\vec{d} = \vec{z} - \vec{o}$$

We then calculate the projection of $\vec{d}$ on eigenvectors of $\mathbf{S}$:

$$\vec{b} = \vec{d}\mathbf{K}$$

Here, $\mathbf{K}$ is the eigenvector matrix of $\mathbf{S}$. The fitness is then calculated as

$$\omega = \exp\left(-\left(\sum_{i=1}^{n} \frac{b_i^2}{2E_i}\right)\right) \tag{5}$$

Here, b_i is the i-th element of $\vec{b}$ and E_i is the n-th eigenvalue of $\mathbf{S}$. If an eigenvalue of $\mathbf{S}$ (e.g., E_i) is zero, the corresponding term in (5) ($\frac{{b_i}^2}{2E_i}$) would be dropped. The biological interpretation of such a situation is

the lack of selection on a specific phenotypic dimension, in which case the phenotypic dimension with no selection should not be considered when calculating fitness.

The specific $\mathbf{S}$-matrix we considered in the simulations, $\mathbf{S}_{10}$ is a 10×10 matrix with all diagonal elements equal to 10 and all off-diagonal elements equal to zero.

Simulation of evolution

For each mutational architecture, we simulated 500 replicate lineages and tracked each lineage's population mean phenotypes through time. The phenotype at the beginning was set to be the origin point, which is also the optimum phenotype, for all simulations.

Simulation along each replicate lineage followed a sequential fixation (i.e., strong-selection-weak-mutation) model [Gillespie, 1983]. The simulation lasted lasted 5×10^4 time steps for each lineage, whereas simulations of evolution under stabilizing selection lasted 2×10^3 time steps for each lineage. In each time step, we first determined the number of mutations that would occur, n_t, which was sampled from a Poisson distribution with mean equal to the total mutation rate U. The probability that the mutation affects a specific trait or module is equal to the ratio of the corresponding trait or module's mutation rate over the total mutation rate. After the trait or module being affected is determined, the mutation's phenotypic effect would be sampled from the corresponding distribution.

After determining the mutation's phenotypic effect, the mutant's fitness was calculated following equation (5). The coefficient of selection was calculated as $s = \omega/\omega_A - 1$, where ω and ω_A are mutant and ancestral fitness, respectively. Given the effective population size N_e, fixation probability of the mutation is calculated as following Kimura [1962]:

$$p_f = \frac{1 - \exp(-2s)}{1 - \exp(-4N_e s)} \tag{6}$$

With probability p_f, phenotypic effect of the mutation would be added to the population's mean phenotype before the next mutation is considered. We had effective population size $N_e = 10^3$ for all simulations.

After the simulations, we calculated the evolutionary variance-covariance matrix ($\mathbf{R}$) and compared its diagonal elements (i.e., the $\mathbf{R}$ variances) to those of $\mathbf{M}$ (i.e., the $\mathbf{M}$ variances). Degree of alignment between

$\mathbf{M}$ and $\mathbf{R}$ is measured by Pearson's correlation coefficient between log-transformed $\mathbf{M}$ and $\mathbf{R}$ variances, as well as the scaling coefficient (i.e., linear regression slope of log-transformed $\mathbf{R}$ variance on log-transformed $\mathbf{M}$ variance). For each trait, we calculated Pearson's correlation coefficient between the amount of variance across lineages and time as a measurement of phylogenetic signal. For the two mutational architectures with modular structures, we examined the scaling between $\mathbf{M}$ and $\mathbf{R}$ variances both across all ten traits and within each module.

All simulations and analyses were done in R.

Results

When pleiotropy is absent, the ten traits had rather similar $\mathbf{R}$ variances, and the $\mathbf{R}$ variances are not significantly correlated with the $\mathbf{M}$ variances (Fig. 1A), indicating the rate of evolutionary divergence is mainly determined by strength of selection rather than mutational input.

Because the per-trait mutation rate and the distribution of mutations' effects on each individual trait remained the same in all the simulations, the proportion of mutations that affect a given trait are more deleterious and less likely to fix as more additional traits pleiotropically are affected by the mutations' pleiotropic effects [McGuigan et al., 2014, Orr, 2000]. Thus, we predicted that $\mathbf{R}$ variances of the traits would decrease as the degree of pleiotropy increased. Indeed, in the presence of universal pleiotropy, $\mathbf{R}$ variances of many traits become much lower than in the absence of pleiotropy (from $\approx 30\%$ to $\approx 1\%$ of the corresponding $\mathbf{R}$ variances without pleiotropy, Fig. 1B). Interestingly, $\mathbf{R}$ variances are strongly correlated to $\mathbf{M}$ ($r > 0.9$ between log-transformed variances, Fig. 1B) and the least-squares regression line of log-transformed $\mathbf{R}$ variance on log-transformed $\mathbf{M}$ variance has a slope < 1 (Fig. 1B), which means the relationship between $\mathbf{M}$ and $\mathbf{R}$ variances are non-linear, and traits with greater $\mathbf{M}$ variance are affected more by constraint resulting from pleiotropy (i.e., ratio of $\mathbf{R}$ variance to $\mathbf{M}$ variance is smaller). In addition, when we examined phylogenetic signal of each trait, which is represented by the correlation between time and phenotypic variance among lineages, we found all traits exhibited stronger phylogenetic signal in the presence of pleiotropy (Table S1).

We also considered mutational architectures where traits are organized into two modules—that is, each mutation affects only one module and affects all traits in this module. When two modules have vastly

different total $\mathbf{M}$ variances, we observed different patterns of $\mathbf{M}$-$\mathbf{R}$ alignment within and across modules: the correlation between $\mathbf{M}$ and $\mathbf{R}$ variances is strong within each module but weak among all ten traits; consistent with this is different scaling of $\mathbf{M}$ and $\mathbf{R}$ in two modules: the module consisting of traits with high $\mathbf{M}$ variances has a much lower regression slope (Fig. 2A). When the two modules have relatively similar total $\mathbf{M}$ variances, in contrast, scaling within the two modules and that across all traits are all similar, characterized with strong correlations and intermediate regression slopes (Fig. 2B). Together, these results suggest hidden modular structures not reflected by the structure of $\mathbf{M}$-matrices can have a profound effect on the alignment of $\mathbf{M}$ and $\mathbf{R}$.

Discussion

In this study, we show a many-to-one mapping from the mutational architecture to the mutational variance-covariance matrix $\mathbf{M}$ and how alternative mutational architectures that map to the same $\mathbf{M}$ can lead to different patterns of macroevolutionary divergence. By simulating evolution in a multi-dimensional phenotypic space, we show the alignment between $\mathbf{M}$ and $\mathbf{R}$ is much stronger when there is universal pleiotropy but no correlation between mutations' effects on different traits compared to that in the absence of pleiotropy. It is worth noting the our observation is similar to patterns previously observed in wing morphology traits in drosophilid flies in terms of evolutionary rates, phylogenetic signal, and the regression slope of $\mathbf{R}$ variances on $\mathbf{M}$ variances after methodological bias is accounted for [Houle et al., 2017, Jiang and Zhang, 2020], suggesting constraint resulting from widespread pleiotropy is a plausible explanation for these observations.

We also examined mutational architectures with modular structures where pleiotropic effects of mutations are restricted within trait modules. Researchers have hypothesized that the underlying genetic architecture of multivariate traits can evolve to be modular such that mutations tend not to affect traits belong to different functional modules simultaneously; this leads to enhanced evolvability and/or mutational robustness [De Visser et al., 2003, Goswami et al., 2014, Melo and Marroig, 2015, Wagner, 1996, Wagner and Altenberg, 1996, Wagner et al., 2007, Wagner and Zhang, 2011]. A widely used method to quantify modularity is the covariance ratio, a statistic that summarizes the overall amount of covariance between traits of different pre-defined modules relative to the overall amount of within-module covariance [Adams, 2016, Adams and Collyer, 2019]. However, as covariances do not fully capture properties the mutational architec-

ture, tests based on the covariance ratio may not capture modularity of the genotype-phenotype map — if pleiotropy is indeed restricted within functional modules, yet mutations' effects on traits are not well correlated, there may not be signal of modularity in covariance-based tests, despite such a scenario shall still be considered a case of modularity. As such hidden modularity can also have significant evolutionary consequences, relying solely on covariance-based tests could also hinder our effort to understand of the evolution of modularity and the influence of modularity on phenotypic evolution.

In this study, we simulated evolution following a sequential fixation model (or strong-selection-weak-mutation model), which corresponds to a scenario where the rate mutations affecting traits of interest is low enough that no more that two alleles segregates in the population at the same time, and focused on the evolution of the population mean phenotype. We chose this as it was more straightforward to study the effects of mutation on alignment under this modeling approach. That being said, the assumptions of the sequential-fixation model are not always valid in nature [McCandlish and Stoltzfus, 2014] and it is an open question as to what extent our results hold in other evolutionary regimes.

The evolutionary problem that motivated this study is the alignment between mutational input and macroevolution, which is also the central focus of this manuscript. The one-to-many mapping between $\mathbf{M}$ and underlying mutational architectures, however, is relevant in a variety of contexts. For instance, it is of great interest whether these alternative mutational architectures lead to different dynamics of adaptive evolution by imposing different levels of constraint. Although the distribution of phenotype at equilibrium is not dependent on details of the mutational architecture as long as there is non-zero mutational variance along dimensions under selection [Sella, 2009], these different mutational architectures could still have a profound impact on transient behaviors of the population—e.g., the amount of time taken to reach the optimum, or the trajectory of movement in the phenotypic space. Evolutionary dynamics during the transient stage is particularly important in the context of evolutionary rescue, where adaptive evolution is required to avoid extinction [Anciaux et al., 2018, Orr and Unckless, 2014].

For the purpose of study, we focused on many-to-one mapping between underlying mutational architectures and $\mathbf{M}$ due to variation in rates and distributions of effect sizes of different classes of mutations. In addition to these aspects of the mutational architecture, there are others that are also not captured by mutational (co)variances yet can potentially affect the pattern of evolutionary divergence at least under some conditions, such as normality of the distribution of phenotypic effects of mutations [Schraiber and Landis,

2015], the number of loci underlying the trait [Lynch, 2020, Lynch and Hagner, 2015, Schraiber and Landis, 2015], and mutational bias [Svensson and Berger, 2019]. Recent studies have also suggested that developmental constraints could be an essential factor that shape the structure of $\mathbf{M}$ and $\mathbf{G}$, and understanding development greatly improve our understanding of the connection between evolution in different timescales [Machado et al., 2023, Milocco and Salazar-Ciudad, 2022, Mongle et al., 2022]. We would encourage theorists and empiricists to pay close attention to the (often implicit) assumptions about underlying architecture of traits embedded in both models and statistical tests.

Conclusions

Our results demonstrate that the $\mathbf{M}$-matrix, a summary statistic commonly used to describe mutational input for quantitative traits, does not fully capture key features of the mutational architecture even when mutations' effects are all additive. We show difference in properties of these mutational architectures can result in difference in the alignment between patterns of mutational input and macroevolutionary divergence. Specifically, if a given $\mathbf{M}$ emerges from a model with high-levels of pleiotropy, we expect a tighter alignment between $\mathbf{M}$ and $\mathbf{R}$ than if the same $\mathbf{M}$ emerged from a model with less pleiotropy but equal per-trait mutation rates. We suggest the many-to-one mapping between the underlying mutational architectures and the $\mathbf{M}$ matrix may contribute to the challenge of explaining macroevolutionary divergence in light of population-level measures.

Acknowledgements

We thank Josh Schraiber, Alex Cope, Arlin Stoltzfus, David McCandlish, Joanna Masel, and two anonymous reviewers for thoughtful comments on the paper that greatly improved this work.. M.P. was supported by startup funds from the University of Southern California.

Code availability

All R code to reproduce our results are available on GitHub (`https://github.com/phylo-lab-usc/m-matrix/`).

References

Ackermann, R. R. and J. M. Cheverud. 2004. Detecting genetic drift versus selection in human evolution. Proceedings of the National Academy of Sciences 101:17946–17951.

Adams, D. C. 2016. Evaluating modularity in morphometric data: challenges with the rv coefficient and a new test measure. Methods in Ecology and Evolution 7:565–572.

Adams, D. C. and M. L. Collyer. 2019. Comparing the strength of modular signal, and evaluating alternative modular hypotheses, using covariance ratio effect sizes with morphometric data. Evolution 73:2352–2367.

Anciaux, Y., L.-M. Chevin, O. Ronce, and G. Martin. 2018. Evolutionary rescue over a fitness landscape. Genetics 209:265–279.

Arnold, S. J. 1992. Constraints on phenotypic evolution. The American Naturalist 140:S85–S107.

Arnold, S. J., R. Bürger, P. A. Hohenlohe, B. C. Ajie, and A. G. Jones. 2008. Understanding the evolution and stability of the g-matrix. Evolution 62:2451–2461.

Arnold, S. J., M. E. Pfrender, and A. G. Jones. 2001. The adaptive landscape as a conceptual bridge between micro-and macroevolution. Microevolution rate, pattern, process Pages 9–32.

Bolstad, G. H., T. F. Hansen, C. Pélabon, M. Falahati-Anbaran, R. Pérez-Barrales, and W. S. Armbruster. 2014. Genetic constraints predict evolutionary divergence in dalechampia blossoms. Philosophical Transactions of the Royal Society B: Biological Sciences 369:20130255.

Boyle, E. A., Y. I. Li, and J. K. Pritchard. 2017. An expanded view of complex traits: from polygenic to omnigenic. Cell 169:1177–1186.

Chakraborty, R. and M. Nei. 1982. Genetic differentiation of quantitative characters between populations or species: I. mutation and random genetic drift. Genetics Research 39:303–314.

Cheverud, J. M. 1984. Quantitative genetics and developmental constraints on evolution by selection. Journal of theoretical biology 110:155–171.

Chevin, L.-M., G. Martin, and T. Lenormand. 2010. Fisher's model and the genomics of adaptation: restricted pleiotropy, heterogenous mutation, and parallel evolution. Evolution 64:3213–3231.

De Visser, J. A. G., J. Hermisson, G. P. Wagner, L. A. Meyers, H. Bagheri-Chaichian, J. L. Blanchard, L. Chao, J. M. Cheverud, S. F. Elena, W. Fontana, et al. 2003. Perspective: evolution and detection of genetic robustness. Evolution 57:1959–1972.

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

Gillespie, J. H. 1983. Some properties of finite populations experiencing strong selection and weak mutation. The American Naturalist 121:691–708.

Goswami, A., J. B. Smaers, C. Soligo, and P. D. Polly. 2014. The macroevolutionary consequences of phenotypic integration: from development to deep time. Philosophical Transactions of the Royal Society B: Biological Sciences 369:20130254.

Gould, S. J. 1980. Is a new and general theory of evolution emerging? Paleobiology 6:119–130.

Halligan, D. L. and P. D. Keightley. 2009. Spontaneous mutation accumulation studies in evolutionary genetics. Annual Review of Ecology, Evolution, and Systematics 40:151–172.

Hansen, T. F. 2006. The evolution of genetic architecture. Annu. Rev. Ecol. Evol. Syst. 37:123–157.

Hansen, T. F. and D. Houle. 2004. Evolvability, stabilizing selection, and the problem. Phenotypic integration: Studying the ecology and evolution of complex phenotypes Page 130.

Harmon, L. J., M. W. Pennell, L. F. Henao-Diaz, J. Rolland, B. N. Sipley, and J. C. Uyeda. 2021. Causes and consequences of apparent timescaling across all estimated evolutionary rates. Annual Review of Ecology, Evolution, and Systematics 52:587–609.

Hill, W. G. and X.-S. Zhang. 2012a. Assessing pleiotropy and its evolutionary consequences: pleiotropy is not necessarily limited, nor need it hinder the evolution of complexity. Nature Reviews Genetics 13:296–296.

Hill, W. G. and X.-S. Zhang. 2012b. On the pleiotropic structure of the genotype–phenotype map and the evolvability of complex organisms. Genetics 190:1131–1137.

Ho, W.-C. and J. Zhang. 2014. The genotype–phenotype map of yeast complex traits: basic parameters and the role of natural selection. Molecular biology and evolution 31:1568–1580.

Houle, D., G. H. Bolstad, and T. F. Hansen. 2020. Fly wing evolutionary rate is a near-isometric function of mutational variation. bioRxiv Pages 2020–08.

Houle, D., G. H. Bolstad, K. van der Linde, and T. F. Hansen. 2017. Mutation predicts 40 million years of fly wing evolution. Nature 548:447–450.

Jiang, D. and J. Zhang. 2020. Fly wing evolution explained by a neutral model with mutational pleiotropy. Evolution 74:2158–2167.

Jones, A. G., S. J. Arnold, and R. Bürger. 2003. Stability of the g-matrix in a population experiencing pleiotropic mutation, stabilizing selection, and genetic drift. Evolution 57:1747–1760.

Jones, A. G., S. J. Arnold, and R. Bürger. 2007. The mutation matrix and the evolution of evolvability. Evolution 61:727–745.

Jones, A. G., R. Bürger, and S. J. Arnold. 2014. Epistasis and natural selection shape the mutational architecture of complex traits. Nature communications 5:3709.

Kimura, M. 1962. On the probability of fixation of mutant genes in a population. Genetics 47:713.

Lande, R. 1976. Natural selection and random genetic drift in phenotypic evolution. Evolution Pages 314–334.

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

Lande, R. 1980. The genetic covariance between characters maintained by pleiotropic mutations. Genetics 94:203–215.

Liu, X., Y. I. Li, and J. K. Pritchard. 2019. Trans effects on gene expression can drive omnigenic inheritance. Cell 177:1022–1034.

Lynch, M. 2020. The evolutionary scaling of cellular traits imposed by the drift barrier. Proceedings of the National Academy of Sciences 117:10435–10444.

Lynch, M. and K. Hagner. 2015. Evolutionary meandering of intermolecular interactions along the drift barrier. Proceedings of the National Academy of Sciences 112:E30–E38.

Lynch, M. and W. G. Hill. 1986. Phenotypic evolution by neutral mutation. Evolution 40:915–935.

Machado, F. A., C. S. Mongle, G. Slater, A. Penna, A. Wisniewski, A. Soffin, V. Dutra, and J. C. Uyeda. 2023. Rules of teeth development align microevolution with macroevolution in extant and extinct primates. Nature Ecology & Evolution Pages 1–11.

McCandlish, D. M. and A. Stoltzfus. 2014. Modeling evolution using the probability of fixation: history and implications. The Quarterly review of biology 89:225–252.

McGlothlin, J. W., M. E. Kobiela, H. V. Wright, D. L. Mahler, J. J. Kolbe, J. B. Losos, and E. D. Brodie III. 2018. Adaptive radiation along a deeply conserved genetic line of least resistance in anolis lizards. Evolution Letters 2:310–322.

McGuigan, K. 2006. Studying phenotypic evolution using multivariate quantitative genetics. Molecular ecology 15:883–896.

McGuigan, K., J. M. Collet, S. L. Allen, S. F. Chenoweth, and M. W. Blows. 2014. Pleiotropic mutations are subject to strong stabilizing selection. Genetics 197:1051–1062.

Melo, D. and G. Marroig. 2015. Directional selection can drive the evolution of modularity in complex traits. Proceedings of the National Academy of Sciences 112:470–475.

Milocco, L. and I. Salazar-Ciudad. 2022. Evolution of the g matrix under nonlinear genotype-phenotype maps. The American Naturalist 199:420–435.

Mongle, C. S., A. Nesbitt, F. A. Machado, J. B. Smaers, A. H. Turner, F. E. Grine, and J. C. Uyeda. 2022. A common mechanism drives the alignment between the micro-and macroevolution of primate molars. Evolution 76:2975–2985.

Nei, M. 2013. Mutation-driven evolution. OUP Oxford.

Opedal, Ø. H., W. S. Armbruster, T. F. Hansen, A. Holstad, C. Pélabon, S. Andersson, D. R. Campbell, C. M. Caruso, L. F. Delph, C. G. Eckert, et al. 2023. Evolvability and trait function predict phenotypic divergence of plant populations. Proceedings of the National Academy of Sciences 120:e2203228120.

Orr, H. A. 2000. Adaptation and the cost of complexity. Evolution 54:13–20.

Orr, H. A. and R. L. Unckless. 2014. The population genetics of evolutionary rescue. PLoS genetics 10:e1004551.

Paaby, A. B. and M. V. Rockman. 2013. The many faces of pleiotropy. Trends in genetics 29:66–73.

Porto, A., H. Sebastião, S. E. Pavan, J. L. VandeBerg, G. Marroig, and J. M. Cheverud. 2015. Rate of evolutionary change in cranial morphology of the marsupial genus monodelphis is constrained by the availability of additive genetic variation. Journal of evolutionary biology 28:973–985.

Rennison, D. J. and C. L. Peichel. 2022. Pleiotropy facilitates parallel adaptation in sticklebacks. Molecular ecology 31:1476–1486.

Rolland, J., L. F. Henao-Diaz, M. Doebeli, R. Germain, L. J. Harmon, L. L. Knowles, L. H. Liow, J. E. Mank, A. Machac, S. P. Otto, et al. 2023. Conceptual and empirical bridges between micro-and macroevolution. Nature Ecology & Evolution Pages 1–13.

Schluter, D. 1996. Adaptive radiation along genetic lines of least resistance. Evolution 50:1766–1774.

Schraiber, J. G. and M. J. Landis. 2015. Sensitivity of quantitative traits to mutational effects and number of loci. Theoretical population biology 102:85–93.

Sella, G. 2009. An exact steady state solution of fisher's geometric model and other models. Theoretical population biology 75:30–34.

Stoltzfus, A. 2021. Mutation, randomness, and evolution. Oxford University Press.

Svensson, E. I. 2022. Multivariate selection and the making and breaking of mutational pleiotropy. Evolutionary Ecology 36:807–828.

Svensson, E. I., S. J. Arnold, R. Bürger, K. Csilléry, J. Draghi, J. M. Henshaw, A. G. Jones, S. De Lisle, D. A. Marques, K. McGuigan, et al. 2021. Correlational selection in the age of genomics. Nature ecology & evolution 5:562–573.

Svensson, E. I. and D. Berger. 2019. The role of mutation bias in adaptive evolution. Trends in ecology & evolution 34:422–434.

Tenaillon, O. 2014. The utility of fisher's geometric model in evolutionary genetics. Annual review of ecology, evolution, and systematics 45:179–201.

Wagner, G. P. 1996. Homologues, natural kinds and the evolution of modularity. American Zoologist 36:36–43.

Wagner, G. P. and L. Altenberg. 1996. Perspective: complex adaptations and the evolution of evolvability. Evolution 50:967–976.

Wagner, G. P., J. P. Kenney-Hunt, M. Pavlicev, J. R. Peck, D. Waxman, and J. M. Cheverud. 2008. Pleiotropic scaling of gene effects and the 'cost of complexity'. Nature 452:470–472.

Wagner, G. P., M. Pavlicev, and J. M. Cheverud. 2007. The road to modularity. Nature Reviews Genetics 8:921–931.

Wagner, G. P. and J. Zhang. 2011. The pleiotropic structure of the genotype–phenotype map: the evolvability of complex organisms. Nature Reviews Genetics 12:204–213.

Walsh, B. and M. W. Blows. 2009. Abundant genetic variation+ strong selection= multivariate genetic constraints: a geometric view of adaptation. Annual Review of Ecology, Evolution, and Systematics 40:41–59.

Wang, Z., B.-Y. Liao, and J. Zhang. 2010. Genomic patterns of pleiotropy and the evolution of complexity. Proceedings of the National Academy of Sciences 107:18034–18039.

Welch, J. J. and D. Waxman. 2003. Modularity and the cost of complexity. Evolution 57:1723–1734.

Zhang, J. 2023. Patterns and evolutionary consequences of pleiotropy. Annual Review of Ecology, Evolution, and Systematics 54:1–19.

Tables

Table 1: Definition of simulation parameters.

Parameter	Definition
z_1, z_2	Phenotypic trait 1 and trait 2
$\mathbf{M}$	Mutational variance-covariance matrix, $\mathbf{M}$-matrix
$\mathbf{M}_P$	Mutational variance-covariance matrix with only pleiotropic mutations considered
$\mathbf{G}$	Genetic variance-covariance matrix, $\mathbf{G}$-matrix
$\mathbf{R}$	Evolutionary variance-covariance matrix, $\mathbf{R}$-matrix
$\mathbb{P}$	Vector of mutational parameters, $\{U_1, \sigma_{e.1}, U_2, \sigma_{e.2}, U_P, \rho, \sigma_{p.1}, \sigma_{p.2}\}$
U_1	Rate of mutations that exclusively affect z_1
U_2	Rate of mutations that exclusively affect z_2
U_P	Rate of pleiotropic mutations that affect both z_1 and z_2
ρ	Correlation between pleiotropic mutations' effects on z_1 and z_1
$\sigma_{e.1}$	Standard deviation of effects of mutations that exclusively affect z_1
$\sigma_{e.2}$	Standard deviation of effects of mutations that exclusively affect z_2
$\sigma_{p.1}$	Standard deviation of pleiotropic mutations' effects on z_1
$\sigma_{p.2}$	Standard deviation of pleiotropic mutations' effects on z_2
C_P, C_ρ, $C_{p.1}$, $C_{p.2}$	Rescaling coefficients multiplied to U_P, ρ, $\sigma_{p.1}$, and $\sigma_{p.2}$
C_1, C_2	Rescaling coefficients multiplied to $U_1\sigma_1^2$ and $U_2\sigma_2^2$
$\mathbf{S}$	Matrix describing the fitness function
$\mathbf{K}$	Eigenvector matrix
ω	Fitness
ω_A	Ancestral fitness
s	Coefficient of selection of a mutation
N_e	Effective population size

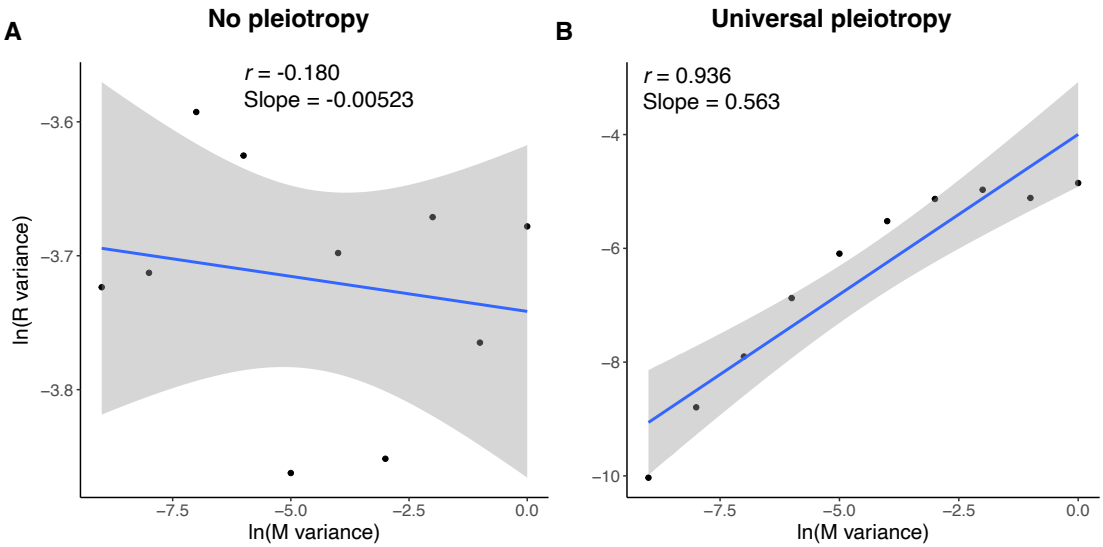

Figure 1: Correlation between log-transformed $\mathbf{M}$ and $\mathbf{R}$ variances of 10 traits when there is no pleiotropy (A) and when there is universal pleiotropy (B). The blue line in each panel is the least-squares regression line, and the gray area is the 95% confidence interval of the regression line.

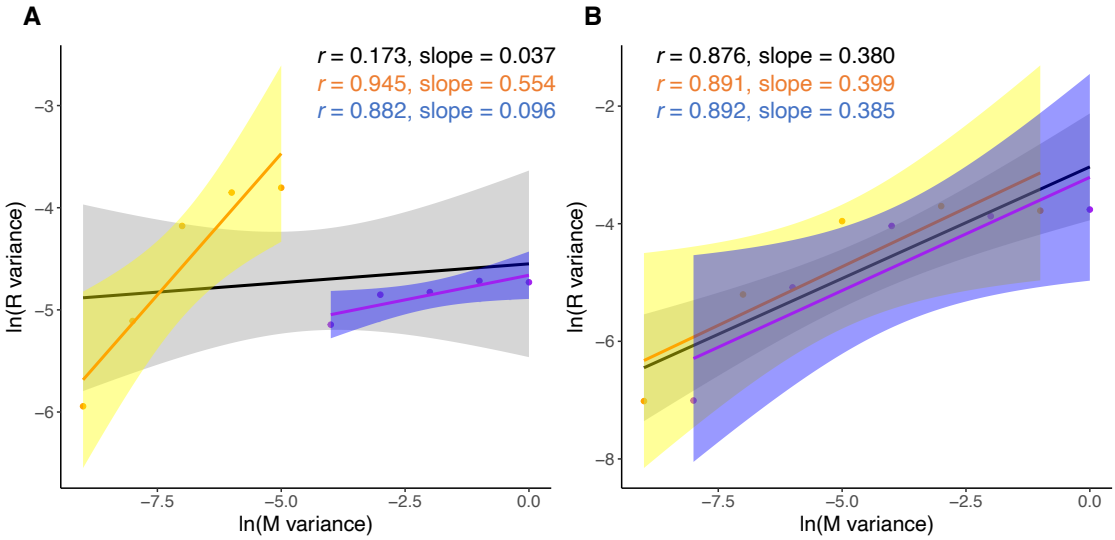

Figure 2: Correlation between log-transformed $\mathbf{M}$ and $\mathbf{R}$ variances of 10 traits when there traits are organized into two modules (e.g., each mutation affects only one module, and affects all traits within the module being affected). (A) five traits with lowest $\mathbf{M}$ variances (i.e., traits 1-5) are in one module, while the other five (i.e., traits 6-10) are in the other. (B) Traits 1, 3, 5, 7, and 9 are in one module, while traits 2, 4, 6, 8, and 10 are in the other module. The black line in each panel is the least-squares regression line across all 10 traits, and the gray area is the 95% confidence interval of the regression line. The purple line in each panel is the within-module regression line for the module with greater total $\mathbf{M}$ variance, with the blue shadowed area being the 95% confidence interval; the orange line in each panel is the within-module regression line for the module with smaller total $\mathbf{M}$ variance, with the yellow shadowed area being the 95% confidence interval.

Supplementary Materials

Table S1: Phylogenetic signal, measured as Pearson's correlation between phenotypic variance among lineages and time, of traits in a 10-dimensional phenotypic space.

Trait #	**Mutational variance**	**Phylogenetic signal**	
		No pleiotropy	Universal pleiotropy
1	exp(-9)	0.773	0.990
2	exp(-8)	0.787	0.989
3	exp(-7)	0.785	0.996
4	exp(-6)	0.822	0.982
5	exp(-5)	0.715	0.979
6	exp(-4)	0.761	0.994
7	exp(-3)	0.712	0.991
8	exp(-2)	0.751	0.986
9	exp(-1)	0.689	0.978
10	1	0.677	0.986