A pleiotropic hitchhiking model recapitulates alignments between fly wing divergence and variation

Haoran Cai*

Department of Ecology and Evolutionary Biology, UCLA.
610 Charles E Young Drive
Los Angeles, CA 90095-1606, USA

ARTICLE TYPE: Research Article

LIST OF ELEMENTS: Manuscript, color version of figures 1-6

ONLINE ENHANCEMENTS: Online Supplementary Material

KEYWORDS: *Drosophila* Wing shape; Rate paradox; Micro-macro evolutionary alignment; Mutational Pleiotropy

*To whom correspondence should be addressed. hrcai@ucla.edu

Abstract

The alignment between mutational variance, standing genetic variance, and macroevolutionary divergence in *Drosophila* wing shape presents a rate paradox: evolution follows mutational lines of least resistance, yet proceeds orders of magnitude slower than quantitative genetic theory predicts. Two explanations have been proposed: "deleterious pleiotropy," where stabilizing selection on unmeasured traits outside the wing morphology complex constrains evolution, and "correlational selection," where selection acts directly on trait covariances. However, recent empirical work has found little evidence of fitness costs associated with wing shape variation beyond effects on flight performance, undermining the deleterious pleiotropy hypothesis. Here, I propose and evaluate a pleiotropic hitchhiking model, in which natural selection targets a single functional trait (wing size) while the complex geometry of wing veins evolves as a correlated response shaped by the structure of mutational variance. Using individual-based simulations parameterized with empirical mutational covariances, I show that univariate selection on a primary trait can reproduce the observed alignment among mutational variance, genetic variance, and divergence. Importantly, the pleiotropic hitchhiking model can also generate slower-than-neutral divergence rates, offering a resolution to the rate paradox without invoking hidden pleiotropic costs on unmeasured traits.

Introduction

A persistent question in evolutionary biology concerns the relationship between microevolution and macroevolution: can macroevolutionary patterns be explained by microevolutionary processes, and conversely, can population-level changes be predicted from macroevolutionary trends (Rolland et al. 2023, Pavlicev 2025, Holstad et al. 2024, Cai et al. 2024, Saito et al. 2025, Melo et al. 2016)? Microevolution refers to the genetic and phenotypic changes within populations driven by mutation, selection, migration, and drift. Macroevolution describes the broader dynamics of life, including the origination and extinction of lineages and large-scale patterns of phenotypic divergence. While attempts to quantitatively link these two scales have yielded mixed success, a comprehensive framework explaining these variable outcomes remains elusive (Schluter 2024).

A notable success in aligning these scales comes from the evolution of wing shape in Drosophila and other Diptera. In this system, a strong positive correlation is consistently observed between the structure of mutational variance (the M-matrix), standing additive genetic variance (the G-matrix), and rates of macroevolutionary divergence (the R-matrix) (Houle et al. 2017). This alignment across different evolutionary timescales offers compelling evidence that non-random phenotypic variation generated by mutation is profoundly linked to long-term evolutionary outcomes.

This alignment, however, presents a paradox. A straightforward explanation is that standing variation limits the rate of evolution (Schluter 1996, Houle et al. 2017). Yet, this hypothesis is challenged by the observation that evolutionary rates are orders of magnitude slower than what the abundant genetic variation would theoretically permit (Houle et al. 2017, Rohner and Berger 2023). This discrepancy reveals a key puzzle: while the direction of evolution aligns with available variation, the rate of evolution does not, which is inconsistent with simple constraint hypotheses.

One proposed resolution is that most mutational variation is effectively unusable due to deleterious pleiotropic effects on unmeasured traits (Houle et al. 2017). This 'pleiotropic constraint' hypothesis predicts genetic covariation between wing shape and fitness components that are functionally unrelated to wing shape or flight. However, a direct test of this in *S. punctum* found no evidence for such covariances (Rohner and Berger 2025).

Two alternative hypotheses (not mutually exclusive) have been proposed without invoking deleterious pleiotropy:

1. **Moving corridor model** (also known as 'selective-lines-of-least-resistance'): This model posits that selection is multivariate, but that the environment creates selective pressures that align with the organism's variational properties (e.g., the M matrix). The multivariate fitness optimum, θ, can be imagined as a mountain peak, and this peak wanders over geological time due to changing environments. The key assumption is the path of this movement: the vector describing the change in the optimum's position, $\frac{d\theta}{dt}$, does not move randomly through the trait space. Instead, it tends to move along "corridors" or "selective lines of least resistance" that are parallel to the primary axis of mutational variance (M_{max}) (Pennell and Jiang 2024, Arnold 2023).

2. **Correlational selection as the ultimate driver**: Optimal scalings exist among wing shape traits, which then shape both the mutation variance-covariance matrix (M) and genetic variance-covariance matrix (G). Naturally, divergence would align with such optimal scalings; therefore, the alignments between divergence and variation are expected under this scenario.

Here, I propose a pleiotropic hitchhiking model as a parsimonious alternative to these hypotheses. The model posits that natural selection acts primarily on wing size, while the complex geometry of wing veins evolves as a correlated byproduct through mutational pleiotropy. By exploring the pure hitchhiking extreme, this framework asks whether complex multivariate selection is strictly necessary to produce the observed patterns. Using individual-based simulations parameterized with empirical mutational variance-covariance matrices, I show that the pleiotropic hitchhiking model can reproduce the observed M-G-R alignment while also contributing to resolve the rate paradox that has challenged simpler constraint hypotheses.

Results

The Pleiotropic Hitchhiking Model

Existing explanations for the alignment between mutational variance (M), genetic variance (G), and macroevolutionary divergence (R) invoke either deleterious pleiotropy on unmeasured traits or pervasive correlational selection maintaining optimal trait combinations. Here, I propose a simpler alternative: the pleiotropic hitchhiking model. This framework requires neither hidden fitness costs nor complex multivariate selection regimes.

The model is defined as follows. Consider a vector of n wing traits, $z = [z_1, z_2, \ldots, z_n]^T$, where z_1 represents the primary trait under direct selection (wing size) and $z_2, \ldots, z_n$ represent wing

shape coordinates describing vein positioning. The directional selection gradient is:

$$\beta = [\beta_s, \epsilon_2, \epsilon_3, \ldots, \epsilon_n]^T \approx [\beta_s, 0, 0, \ldots, 0]^T$$

where β_s reflects selection on the primary trait (wing size) and the selection gradients for all other traits (ϵ_i) are negligibly small. Under this formulation, wing shape traits are effectively neutral with respect to fitness, but experience indirect evolutionary change through their mutational covariance with the selected trait—the "mutational drag" effect.

A critical prediction of this model is that, given stable M and G matrices over macroevolutionary time, the divergence matrix R should align with both M and G. This alignment emerges not necessarily from correlational selection constraining shape evolution, but from the persistent structure of mutational input channeling neutral trait divergence along predictable directions.

Pleiotropic Hitchhiking Recapitulates the Empirical Alignment Between Divergence and Variation

To test whether such a pleiotropic hitchhiking model can reproduce the empirical alignment patterns observed in fly wings, I conducted individual-based simulations using the empirical mutational variance-covariance matrix M from Houle et al. (2017). Wing size (centroid size) was designated as the primary trait under stabilizing selection on a moving optimum, consistent with substantial evidence that wing size is a direct determinant of aerodynamic performance. Basic physical principles dictate that lift generation scales with wing surface area, while power requirements for flight scale differently with size. This creates strong selection on the ratio of wing size to body mass across diverse ecological contexts (Yenisetti and Hegde 2003, Kotyk and Varadínová 2017, Altizer and Davis 2010).

The simulation protocol was as follows: 20 replicate populations were initialized from an ancestral population experiencing a burn-in session. To allow populations to diverge, each replicate was assigned an optimum moving rate drawn from a normal distribution with mean zero and standard deviation σ. This means the optimum translates at a constant velocity for a given lineage, $\theta_i(t) = v_i t$, effectively generating a sustained directional selection scenario rather than a Brownian motion random walk. This design enables populations to track different selective targets at different speeds, generating among-lineage divergence in the primary trait (size). Larger σ values produce faster and more variable divergence among replicates, while smaller σ values constrain populations to remain closer to the ancestral optimum. The divergence matrix R

was computed from the among-population variance in trait means, while the within-population genetic variance-covariance matrix G was recorded at the final generation and averaged across 20 replicate populations.

To quantify the alignment between matrices, I employed Krzanowski's common subspace analysis following previous approaches (Houle et al. 2020, Jiang and Zhang 2020). This method projects the matrices onto a common set of orthogonal axes (e.g., the eigenvectors of M) and compares the variance along each axis on a logarithmic scale. A strong linear relationship indicates proportional scaling of variances—the signature of alignment between matrices.

The results show that the pleiotropic hitchhiking model successfully reproduces the empirical patterns (Fig. 1). When comparing $\log_{10}$-transformed variances along the principal axes of M, both R and G showed strong positive correlations with M. This alignment was robust across two orders of magnitude variation in selection strength on the primary trait (Fig. S6) and across different reference matrices (Fig. S7). These findings indicate that a model of univariate selection on wing size, combined with stable mutational covariances, is sufficient to generate the observed M-G-R alignment without invoking correlational selection on shape.

Reduced Rate of Shape Divergence Under Pleiotropic Hitchhiking

The alignment between M, G, and R alone does not distinguish pleiotropic hitchhiking from neutral evolution: under complete neutrality, divergence would also scale proportionally with mutational variance, yielding a slope of 1 in the matrix comparison analyses (Houle et al. 2017). The critical observation that requires explanation is that wing shape evolves orders of magnitude more slowly than expected from the abundant genetic variation available.

The pleiotropic hitchhiking model offers a resolution to this rate paradox. Although shape traits are not directly selected, they experience apparent stabilizing selection through their mutational covariance with wing size. Mutations affecting shape also tend to affect size, and those with large effects on the selectively constrained trait are eliminated. The net result is an effective constraint on shape evolution despite the absence of direct selection on shape.

Using a simplified two-trait model (Appendix A; Figs. S1–S3), I show that stronger selection on the primary trait constrains the drift of correlated neutral traits through two reinforcing mechanisms: (i) apparent stabilizing selection erodes the genetic variance of the neutral trait (G_{BB}), reducing the total variance available for drift; and (ii) stronger selection drives the genetic correlation (ρ_G) toward the mutational correlation, so that a larger fraction of the remaining

variance is "leashed" to the selected trait and cannot drift independently. A third effect acts in the opposite direction: strong selection reduces the effective population size (N_e), which increases drift (Singh and Agrawal 2022, Taylor 2009). Nevertheless, the drift-constraining effects of variance reduction and correlation constraint dominate, and together they produce a substantial net reduction in the effective drift rate of the neutral trait (see Appendix A).

Simulation results using the full empirical M matrix support these predictions. Figure 2 shows the effective drift rate for all 24 wing shape traits under pleiotropic hitchhiking. Under stronger stabilizing selection on wing size ($V_s = 0.05$), the drift rate of every shape trait is substantially reduced (by approximately two orders of magnitude) compared to both weak selection ($V_s = 0.5$) and the completely neutral scenario, indicating that selection on a single primary trait can constrain the drift of correlated neutral traits. Analogous patterns emerge in the two-trait toy model, where varying selection strength systematically modulates the effective drift rate (Fig. S3). Moreover, the relationship between mutational variance and drift rate across traits (Fig. 3) shows that traits with higher mutational input diverge faster, preserving the rate–variance alignment even under selection.

Selective Depletion of Genetic Variance in Size

Beyond reducing drift rates, pleiotropic hitchhiking makes a further testable prediction: stabilizing selection on the primary trait (wing size) should deplete its standing genetic variance relative to mutational input more than it depletes genetic variance of the neutral shape traits. To quantify this, I examined the ratio G_{ii}/M_{ii} for each trait i, which reflects how much standing genetic variance is maintained at mutation–selection–drift balance per unit of mutational input. Under pure drift, this ratio is expected to be approximately equal for all traits; selection on a given trait should reduce its G/M ratio below those of unselected traits.

Results from individual-based simulations support this prediction (Fig. 4). Under neutral evolution, the G/M ratio is approximately uniform across all traits, including size. Under the pleiotropic hitchhiking model, the size trait's G/M ratio is markedly reduced compared to shape traits, consistent with selection selectively eroding genetic variance along the size axis.

To test whether this prediction holds empirically, I computed the per-trait G/M ratio using the empirical G and M matrices estimated for *D. melanogaster* wing traits from Houle et al. (2017). Notably, the G/M ratio for centroid size (natural-log-transformed) is the lowest of all 25 traits, under both the homozygous ($G/M_{\text{hom}} = 72$; shape median $= 347$) and heterozygous ($G/M_{\text{het}} = 144$; shape median $= 842$) mutational variance estimates (Fig. 5). Size retains only

$\sim$21% (M_{hom}) to $\sim$17% (M_{het}) as much standing genetic variance per unit of mutational input compared to the median shape trait.

Hitchhiking and correlational selection leave distinct imprints on G

The pleiotropic hitchhiking model and the correlational selection hypothesis make distinct predictions about the eigenvalue structure of G relative to M. If correlational selection is the primary force shaping genetic covariances—favoring specific trait combinations that mirror the mutational architecture—then the equilibrium G should preserve the eigenvalue structure of M, maintaining similar eccentricity (i.e., similar ratios among eigenvalues). Conversely, under pleiotropic hitchhiking, where only a single primary trait is under selection, variance along the leading eigenvector of M (which is size-dominated) should be preferentially eroded, while variance in orthogonal directions remains largely unconstrained. This predicts that G should be less eccentric than M—a "sphericalization" effect.

Dugand et al. (2021) found that mutational correlations among fly wing traits are consistently stronger than the corresponding additive genetic correlations. Furthermore, mutational variance is highly concentrated along a single multivariate axis (high eccentricity), while genetic variance is more evenly distributed across axes (lower eccentricity). Reanalysis of G and M from Houle et al. (2017) reveals a similar pattern: the ratio of the first eigenvalue is substantially larger for M than for G (Fig. S5 and Appendix B). These independent datasets converge on a consistent finding: selection reduces, rather than reinforces, the eccentricity imposed by mutational input.

To formally test these competing predictions, I conducted simulations comparing equilibrium G under two scenarios (using M obtained from Dugand et al. 2021): (1) the pleiotropic hitchhiking model, where only the primary trait is under stabilizing selection; and (2) correlational selection, where the selection matrix S is proportional to M, imposing multivariate stabilizing selection that favors the trait combinations encoded by mutational pleiotropy. Prior to simulating, the empirical M matrix was forced to be positive semi-definite before being used in `mvrnorm` to generate pleiotropic mutational effects. Three quantitative diagnostics were applied to the simulated equilibrium G matrices (5 replicates per scenario) to formally discriminate the two hypotheses.

The results (Fig. 6, Fig. S10, and Fig. S11) show that the eigenvalue spectrum of equilibrium G under hitchhiking is visibly less eccentric than that of M: the proportion of total genetic variance explained by the leading eigenvalue is reduced under hitchhiking, with the residual variance redistributed among the rest of the eigenvectors; in contrast, correlational selection preserves

the eigenvalue proportions close to those of M. This flattening is also captured by the effective dimensionality (n_{eff}; Kirkpatrick 2009): under hitchhiking, $n_{\text{eff}} = 2.17 \pm 0.35$ at $V_s = 1$ and 2.09 ± 0.15 at $V_s = 5$ (mean $\pm$ s.d. across replicates), both substantially exceeding the M baseline ($n_{\text{eff}} = 1.80$, computed directly from the mutational covariance matrix), whereas correlational selection leaves n_{eff} essentially unchanged (1.79 ± 0.19 at $V_s = 1$; 1.88 ± 0.26 at $V_s = 5$). Second, the directional evolvability ratio—the ratio of evolvability along the size direction to the average evolvability orthogonal to it ($e_{\parallel}/e_{\perp}$; Hansen and Houle 2008)—drops from 1.11 in M (computed directly) to 0.14 ± 0.07 at $V_s = 1$ and 0.31 ± 0.05 at $V_s = 5$ under hitchhiking. Under correlational selection, this ratio is preserved ($e_{\parallel}/e_{\perp} = 1.19 \pm 0.07$ at $V_s = 1$; 1.23 ± 0.20 at $V_s = 5$), because proportional multivariate selection erodes variance in all directions equally. The empirical G matrix from Dugand et al. (2021) exhibits the similar qualitative signatures as the hitchhiking simulations: $n_{\text{eff}} = 3.42$ (exceeding both M and the simulated hitchhiking equilibrium) and $e_{\parallel}/e_{\perp} = 0.77$ (below the isotropic baseline of 1), consistent with selective depletion of size-related genetic variance. These diagnostics collectively suggest that selection on a single trait reshapes the eigenvalue structure of G in a manner qualitatively distinct from multivariate correlational selection. Notably, the empirical $n_{\text{eff}} = 3.42$ exceeds the simulated hitchhiking values at all selection strengths, indicating that additional factors—such as environmental variance, selection on multiple targets, or more complex pleiotropic architectures—contribute to the dimensionality of empirical G beyond what the idealized single-trait model predicts.

These findings should be interpreted cautiously, given that empirically estimated M matrices may carry greater sampling uncertainty due to lower effective sample sizes (Dugand et al. 2021), and the two scenarios tested here represent idealized extremes of a continuum that may include both processes operating simultaneously.

Discussion

The strong alignment between mutational variation (M), genetic variation (G), and species divergence (R) in fly wing shape has long puzzled evolutionary biologists. Here, I have shown that a pleiotropic hitchhiking model—where wing shape evolves as a neutral passenger to stabilizing selection on wing size—can reproduce these empirical patterns without invoking hidden constraints, correlational selection, or complex adaptive landscapes. The model provides a parsimonious explanation for the striking evolutionary conservatism of fly wing allometry, while remaining compatible with the possibility that allometric relationships are optimized by selection (Bolstad et al. 2015, Houle et al. 2019).

The Pleiotropic Hitchhiking Model versus Deleterious Pleiotropy

The pleiotropic hitchhiking model described here differs subtly from the prevailing "deleterious pleiotropy" hypothesis (Houle et al. 2017). The original deleterious pleiotropy model emphasizes horizontal pleiotropy—mutations affecting wing shape are assumed to have deleterious effects on functionally unrelated traits outside the wing trait complex/module (Cai et al. 2023). In contrast, the pleiotropic hitchhiking model invokes vertical pleiotropy within the wing trait complex: mutations affecting size necessarily affect shape — leading to consistent pleiotropy across the genome. This form of pleiotropy is expected given the polygenic nature of shape and size in fly wings (Cai et al. 2023 and Appendix E) and the highly integrated developmental architecture of the wing.

Evidence for Wing Size as the Primary Selection Target

The pleiotropic hitchhiking model requires that wing size experiences stronger direct selection than wing shape. There is considerable evidence supporting this assumption. Wing size is a critical determinant of aerodynamic performance, directly influencing lift generation and power requirements for flight. Research has documented selection on wing size through multiple fitness components: mating success (Yenisetti and Hegde 2003, Kotyk and Varadínová 2017, Blanckenhorn et al. 2021), migration survival (Altizer and Davis 2010, Dockx 2007), thermoregulation (Berwaerts et al. 2001), and fecundity (Chang et al. 2021, McCulloch et al. 2025, Outomuro et al. 2021, Yang et al. 2025). In sepsid flies specifically, a comparative study of sexual selection across seven species found that larger males consistently had higher mating success, and that sexual selection on wing and fore femur shape was closely aligned with static allometry (Blanckenhorn et al. 2021). This alignment is consistent with the hypothesis that apparent selection on shape is a byproduct of selection on overall body size. However, allometry-corrected shape selection often remained significant (Blanckenhorn et al. 2021), indicating that some selection on shape persists independently of size. Because the size-related mating differentials were not corrected for correlated shape variation, the converse interpretation—that selection targets shape and produces apparent selection on size through allometry—cannot be excluded.

Direct evidence for selection on the relative positions of wing vein intersections (i.e., shape independent of size) is scarce. While vein positioning may influence wing flexibility and aerodynamic properties (Combes and Daniel 2003), the functional consequences of natural shape variation appear subtle compared to size effects. In odonates, wing shape does show significant genetic differentiation across latitudinal gradients, with genetic population differences

having two to four times the impact of phenotypic plasticity on shape variation (Johansson et al. 2023). However, even in these systems, wing shape variation is tightly linked to life-history traits through developmental constraints and pleiotropy (Johansson et al. 2023), and body size—which varies clinally with developmental time constraints (Outomuro et al. 2021)—can scale allometrically with wing shape. These findings suggest that size-shape developmental integration provides a general mechanism by which selection on size can indirectly constrain shape evolution, although direct selection on shape cannot be entirely excluded.

However, a critical caveat is that the model's ability to constrain shape evolution depends on the strength of selection on wing size. Under weaker selection ($V_s = 0.5$), the drift rate of shape traits is barely distinguishable from the neutral regime (Fig. 2); only under strong selection ($V_s = 0.05$) does the model generate meaningful rate reductions. In the toy model, sufficiently strong selection on one trait can even produce deterministic-like behavior in the correlated trait (Fig. S3). Therefore, the utility of the pleiotropic hitchhiking model in resolving the rate paradox hinges critically on whether stabilizing selection on wing size is sufficiently strong in nature—an empirical question that warrants direct investigation.

Interpreting the simulated V_s values requires noting that centroid size is natural-log-transformed in the empirical M matrix (Houle et al. 2017): V_s thus measures the width of the fitness function on the log-size scale.

Other Factors Contributing to the Rate Paradox

While the pleiotropic hitchhiking model can generate slower-than-neutral divergence rates, it may not fully account for the magnitude of the rate paradox. Empirical estimates suggest that the observed rate of wing shape divergence is approximately 10^{-4} of the rate expected under neutral evolution (Houle et al. 2017). The reduction in effective drift rate predicted by the pleiotropic hitchhiking model (Fig. 2) may be insufficient to explain such extreme evolutionary stasis.

This discrepancy suggests that additional mechanisms may contribute to constraining wing shape evolution. One possibility is that deleterious horizontal pleiotropy—effects on traits outside the wing shape complex—does play a role, even if such effects were not statistically significant in the Sepsidae system examined by Rohner and Berger (2025). Future work combining the pleiotropic hitchhiking framework with explicit models of additional pleiotropic effects may help reconcile the full magnitude of the rate paradox.

Stability of the Mutational Variance-Covariance Matrix

A key assumption of the pleiotropic hitchhiking model is that the structure of M remains stable over macroevolutionary time. Several lines of evidence support this assumption. First, genetic variance-covariance matrices (G) appear similar between Sepsidae and *Drosophila melanogaster*, which diverged approximately 64 million years ago (Wiegmann et al. 2011, Rohner and Berger 2025). Common subspace analyses comparing G matrices across these taxa reveal substantial alignment regardless of which species is used as the reference (Figs. S8, S9). Given that G reflects both mutational input and selection, conservation of G implies conservation of the underlying M. Second, the highly polygenic genetic architecture of wing shape traits (Appendix E) may enhance the stability of M and G by buffering against the effects of individual mutations (Barton and Turelli 1989). Nevertheless, direct measurement of M across divergent lineages remains an important goal for future research.

Ultimate versus Proximate Causes

It is important to distinguish between proximate and ultimate explanations for the M-G-R alignment. The pleiotropic hitchhiking model provides a proximate mechanism: given a stable M matrix and selection on a single trait, alignment emerges as a natural consequence. However, the ultimate origin of the mutational architecture itself remains an open question. Mutation is evolvable (Jones et al. 2014, Cai et al. 2024), and the structure of M may itself be shaped by selection over deep evolutionary time. Alternatively, the movement of phenotypic optima along directions of maximal mutational variance (the "moving corridor" hypothesis) could align selection with mutational architecture (Houle et al. 2017, Pennell and Jiang 2024). These ultimate-level hypotheses are not mutually exclusive with the pleiotropic hitchhiking model and deserve further investigation.

Materials and Methods

Model Overview

The pleiotropic hitchhiking model, proposed here, posits that natural selection targets only one primary trait, while remaining traits evolve as neutral byproducts through their mutational covariance with the selected trait. The vector of directional selection gradients β is therefore

specified as:

$$\beta = [\beta_s, 0, 0, \ldots, 0]^T$$

where β_s reflects the strength of selection on the primary trait, determined by the distance between the population mean and the current optimum. Under this formulation, evolutionary change in the secondary traits $(z_2, \ldots, z_n)$ occurs solely through mutational covariance with the selected trait—the "mutational drag" effect—rather than through direct selection.

Individual-Based Simulations

I assessed the pleiotropic hitchhiking model using individual-based simulations. I simulated populations of $N = 500$ hermaphroditic, sexually reproducing individuals with non-overlapping generations. Phenotypes were determined by $n = 50$ unlinked diploid loci, with alleles assumed to be fully pleiotropic (i.e., affecting all phenotypic traits). The life cycle proceeded in three discrete steps:

1. **Phenotype and fitness assignment.** Each individual's phenotypic value for every trait was obtained by summing allelic breeding values across all loci, assuming strict additivity (no dominance or epistasis at the phenotypic level, though epistasis may arise at the fitness level). Fitness was then assigned according to a Gaussian stabilizing selection function applied solely to the primary trait:

$$W(z_1) = \exp\left(-\frac{(z_1 - \theta)^2}{2V_s}\right)$$

   where z_1 is the individual's primary trait value, θ is the current phenotypic optimum, and V_s controls the width of the fitness peak (smaller V_s imposes stronger selection).

2. **Selection and mating.** Parents were sampled with replacement in proportion to their fitness, and two distinct individuals were paired to produce each offspring (selfing was excluded). This fitness-weighted sampling was repeated N times per generation to maintain constant population size.

3. **Recombination and mutation.** Offspring genotypes were assembled by independently drawing one allele per locus from each parent, corresponding to free recombination among all loci. Mutations occurred at a rate of $\mu = 0.01$ per allele per generation (genomic mutation rate $U = 0.5$ per individual per generation). Pleiotropic mutational effects

were drawn from a multivariate normal distribution with variance-covariance matrix $\mathbf{M}$, estimated empirically under homozygous conditions by Houle et al. (2017).

Each simulation began with a burn-in period of $10N$ generations, with the optimum set to zero throughout. Following burn-in, 20 replicate populations were evolved for $T = 20000$ generations under divergent directional selection on the primary trait, where each replicate was assigned an optimum moving rate sampled from a normal distribution with mean $\mu = 0$ and standard deviation $\sigma = 0.0001$. Selection strength was varied across simulations by setting $V_s = 0.05$ (strong selection) or $V_s = 0.5$ (weak selection).

Comparing Variance-Covariance Matrices

To quantify the alignment between mutational variance ($\mathbf{M}$), genetic variance ($\mathbf{G}$), and evolutionary divergence ($\mathbf{R}$), I employed the common subspace analysis of Jiang and Zhang (2020). This approach compares matrices along a shared set of orthogonal axes defined by the eigenvectors of $\mathbf{M}$.

Specifically, I decomposed the homozygous mutational variance-covariance matrix ($\mathbf{M}_{hom}$) estimated in *D. melanogaster* into its eigenvector matrix $\mathbf{K}$. For each variance-covariance matrix of interest ($\mathbf{X} \in \{\mathbf{G}, \mathbf{R}\}$), I calculated the variance along each eigenvector as the diagonal entries of $\mathbf{K}^\mathsf{T}\mathbf{X}\mathbf{K}$. I then regressed $\log_{10}$-transformed variances in $\mathbf{R}$ or $\mathbf{G}$ against $\log_{10}$-transformed variances in $\mathbf{M}$ using ordinary least squares (OLS), obtaining regression slopes and coefficients of determination (R^2). To avoid comparisons along dimensions with negligible variance, analyses were restricted to the first 18 eigenvectors of $\mathbf{M}$.

Since this matrix comparison method is known to, in some cases, produce downwardly biased slope estimates (Jiang and Zhang 2020, Houle et al. 2020), an additional reference matrix G (from Houle et al. 2017) was applied.

Acknowledgements

I thank David Houle and Patrick Rohner for helpful discussions.

References

Altizer, S. and Davis, A. K. 2010. Populations of monarch butterflies with different migratory behaviors show divergence in wing morphology. Evolution 64:1018–1028.

Arnold, S. J., 2023. Evolutionary quantitative genetics. Oxford University Press.

Barton, N. H. and Turelli, M. 1989. Evolutionary quantitative genetics: how little do we know. Annual review of genetics 23:337–370.

Berwaerts, K., Van Dyck, H., Vints, E., and Matthysen, E. 2001. Effect of manipulated wing characteristics and basking posture on thermal properties of the butterfly pararge aegeria (l.). Journal of Zoology 255:261–267.

Blanckenhorn, W. U., Baur, J., Roy, J., Puniamoorthy, N., Busso, J. P., Schäfer, M. A., and Rohner, P. T. 2021. Comparative sexual selection in field and laboratory in a guild of sepsid dung flies. Animal Behaviour 175:219–230.

Bolstad, G. H., Cassara, J. A., Márquez, E., Hansen, T. F., van der Linde, K., Houle, D., and Pélabon, C. 2015. Complex constraints on allometry revealed by artificial selection on the wing of drosophila melanogaster. Proceedings of the National Academy of Sciences 112:13284–13289.

Cai, H., Geiler-Samerotte, K., and Des Marais, D. L. 2023. Dissecting genetic correlation through recombinant perturbations: the role of developmental bias. bioRxiv Pages 2023–05.

Cai, H., Melo, D., and Des Marais, D. 2024. Disentangling variational bias: the roles of development, mutation and selection. Trends in Genetics .

Chang, H., Guo, X., Guo, S., Yang, N., and Huang, Y. 2021. Trade-off between flight capability and reproduction in acridoidea (insecta: Orthoptera). Ecology and Evolution 11:16849–16861.

Combes, S. A. and Daniel, T. L. 2003. Flexural stiffness in insect wings i. scaling and the influence of wing venation. Journal of experimental biology 206:2979–2987.

Dockx, C. 2007. Directional and stabilizing selection on wing size and shape in migrant and resident monarch butterflies, danaus plexippus (l.), in cuba. Biological Journal of the Linnean Society 92:605–616.

Dugand, R. J., Aguirre, J. D., Hine, E., Blows, M. W., and McGuigan, K. 2021. The contribution of mutation and selection to multivariate quantitative genetic variance in an outbred population of drosophila serrata. Proceedings of the National Academy of Sciences 118:e2026217118.

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

Holstad, A., Voje, K. L., Opedal, Ø. H., Bolstad, G. H., Bourg, S., Hansen, T. F., and Pélabon, C. 2024. Evolvability predicts macroevolution under fluctuating selection. Science 384:688–693.

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

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

Houle, D., Jones, L. T., Fortune, R., and Sztepanacz, J. L. 2019. Why does allometry evolve so slowly? Integrative and comparative biology 59:1429–1440.

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

Johansson, F., Berger, D., Outomuro, D., Sniegula, S., Tunon, M., Watts, P. C., and Rohner, P. T. 2023. Mixed support for an alignment between phenotypic plasticity and genetic differentiation in damselfly wing shape. Journal of Evolutionary Biology 36:368–380.

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

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

Kotyk, M. and Varadínová, Z. 2017. Wing reduction influences male mating success but not female fitness in cockroaches. Scientific reports 7:2367.

McCulloch, G. A., Foster, B. J., Kroos, G. C., Foster, Y., Ni, S., Vogel, M. F., and Waters, J. M. 2025. Dispersal-fecundity trade-offs in wild insect populations. Journal of evolutionary biology 38:430–436.

Melo, D., Porto, A., Cheverud, J. M., and Marroig, G. 2016. Modularity: genes, development, and evolution. Annual review of ecology, evolution, and systematics 47:463–486.

Outomuro, D., Golab, M. J., Johansson, F., and Sniegula, S. 2021. Body and wing size, but not wing shape, vary along a large-scale latitudinal gradient in a damselfly. Scientific Reports 11:18642.

Pavlicev, M. 2025. Cracking open the blackbox of genotype-phenotype map: Crossing the explanatory gap between micro-and macroevolution. Evolution .

Pennell, M. and Jiang, D., 2024. The macroevolutionary adaptive landscape: more than a metaphor?

Rohner, P. T. and Berger, D. 2023. Developmental bias predicts 60 million years of wing shape evolution. Proceedings of the National Academy of Sciences 120:e2211210120.

Rohner, P. T. and Berger, D. 2025. Macroevolution along developmental lines of least resistance in fly wings. Nature Ecology & Evolution Pages 1–13.

Rolland, J., Henao-Diaz, L. F., Doebeli, M., Germain, R., Harmon, L. J., Knowles, L. L., Liow, L. H., Mank, J. E., Machac, A., Otto, S. P., et al. 2023. Conceptual and empirical bridges between micro-and macroevolution. Nature Ecology & Evolution 7:1181–1193.

Saito, K., Tsuboi, M., and Takahashi, Y. 2025. Conserved wing shape variation across biological scales unveils dialectical relationships between micro-and macroevolution. Communications Biology 8:990.

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

Schluter, D. 2024. Variable success in linking micro-and macroevolution. Evolutionary Journal of the Linnean Society 3:kzae016.

Singh, A. and Agrawal, A. F. 2022. Sex-specific variance in fitness and the efficacy of selection. The American Naturalist 199:587–602.

Taylor, J. E. 2009. The genealogical consequences of fecundity variance polymorphism. Genetics 182:813–837.

Wiegmann, B. M., Trautwein, M. D., Winkler, I. S., Barr, N. B., Kim, J.-W., Lambkin, C., Bertone, M. A., Cassel, B. K., Bayless, K. M., Heimberg, A. M., et al. 2011. Episodic radiations in the fly tree of life. Proceedings of the National Academy of Sciences 108:5690–5695.

Yang, J., Yang, C., Lin, H.-w., Lees, A. C., and Tobias, J. A. 2025. Elevational constraints on flight efficiency shape global gradients in avian wing morphology. Current Biology 35:1890–1900.

Yenisetti, S. C. and Hegde, S. N. 2003. Size-related mating and reproductive success in a drosophilid: Phorticella striata. ZOOLOGICAL STUDIES-TAIPEI- 42:203–210.

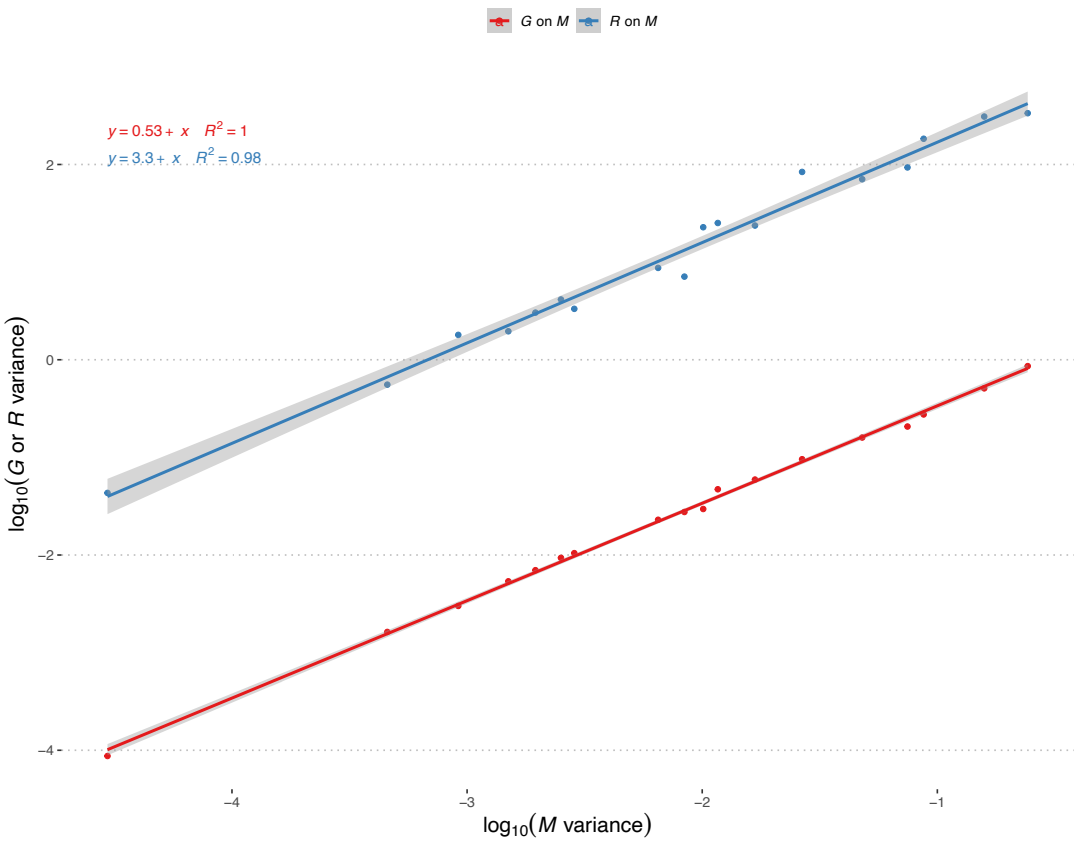

Figure 1: **The pleiotropic hitchhiking model reproduces the empirical alignment between M, G, and R.** Common subspace analysis comparing the divergence matrix R (blue) and genetic variance matrix G (red) to the mutational variance matrix M. Points represent $\log_{10}$ variance along the eigenvectors of M; lines indicate ordinary least squares regression. Strong positive relationships ($R^2 > 0.8$) indicate that both R and G are aligned with the mutational architecture. Only the upper 18 eigenvectors of M were used; wing size was not included in the analyses. Shaded areas show 95% CI. Simulations used $V_s = 0.05$ on wing size (strong selection), $\sigma = 0.0001$, $N = 500$, and 20 replicate populations evolved for 20,000 generations.

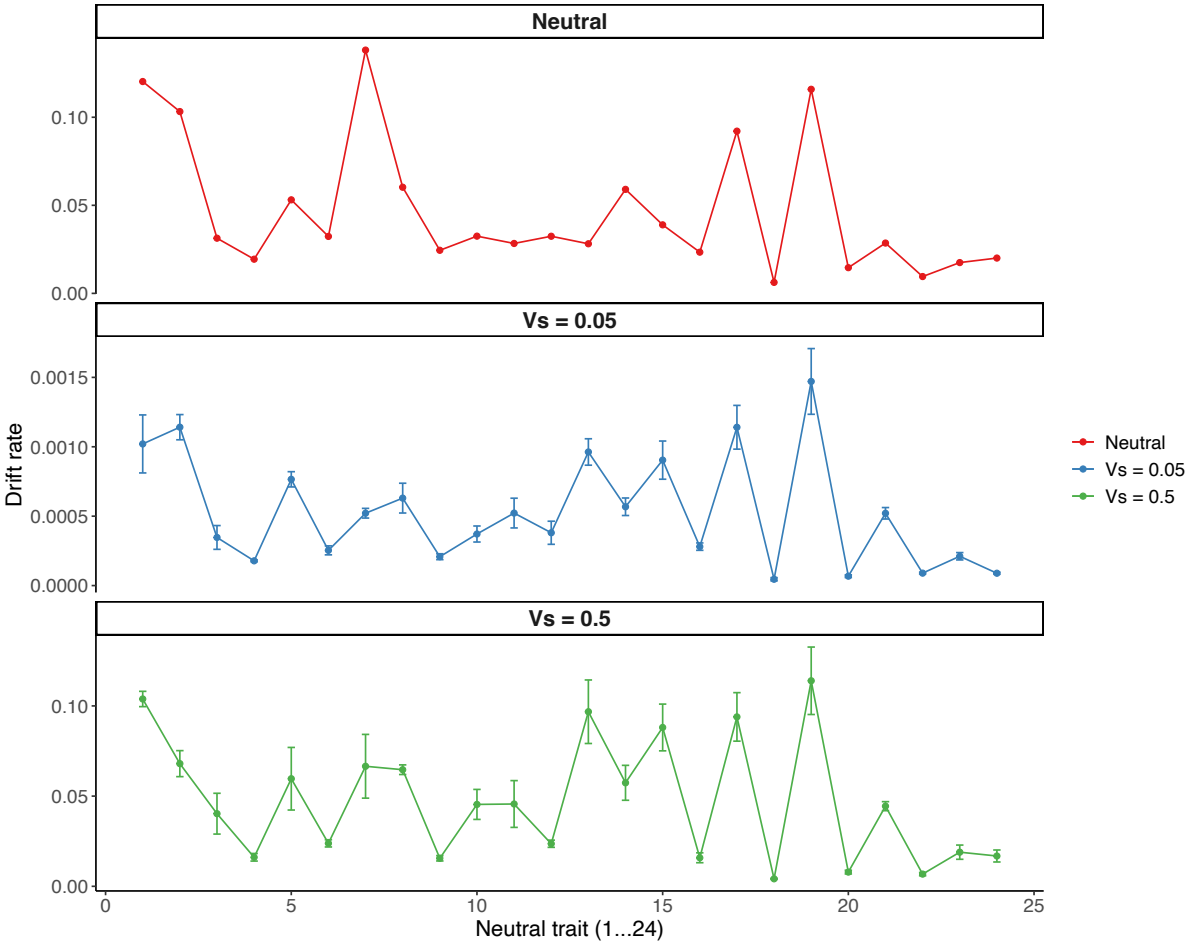

Figure 2: **Effective drift rate for all 24 wing shape traits under pleiotropic hitchhiking.** Each point represents the drift rate (increase in among-population variance per generation) for one wing shape trait. Under strong stabilizing selection on wing size ($V_s = 0.05$, blue), drift rates are substantially reduced compared to weak selection ($V_s = 0.5$, green) and the completely neutral scenario (red, where both wing size and shape are neutral). This indicates that selection on a single primary trait can constrain the drift of correlated neutral traits. Simulations used $\sigma = 0.0001$, $N = 500$, and 20 replicate populations evolved for 20,000 generations. Three independent simulation replicates were used for each V_s condition ($V_s = 0.05$ and $V_s = 0.5$).

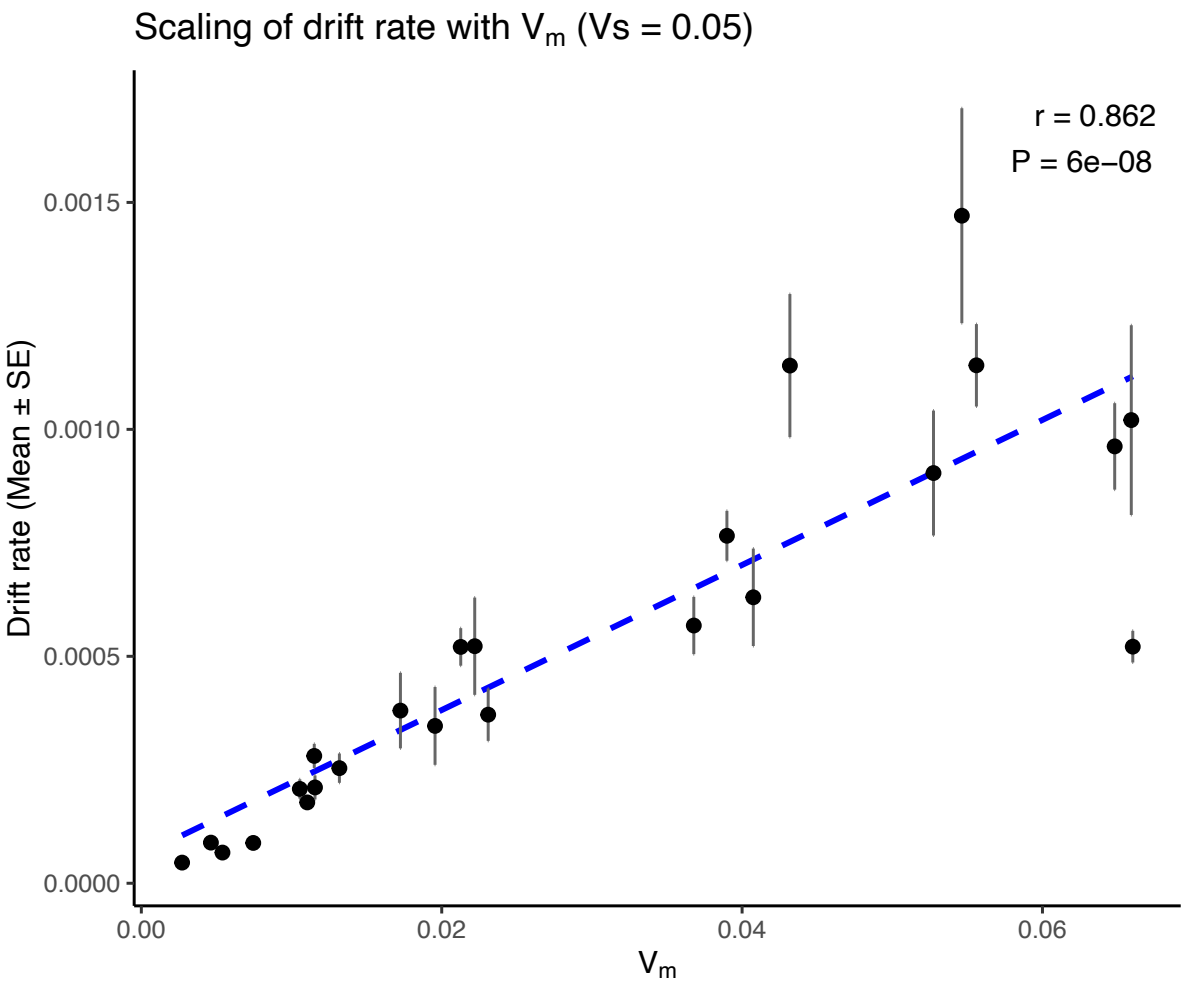

Figure 3: **Drift rate scales with mutational variance under pleiotropic hitchhiking.** Each point represents one trait (wing size plus 24 shape traits). The positive relationship between mutational variance (V_M) and drift rate shows that traits with higher mutational input diverge faster, even when most traits are not under direct selection. This scaling underlies the alignment between M, G, and R. Simulations used $V_s = 0.05$ on wing size, $\sigma = 0.0001$, $N = 500$, and 20 replicate populations. Three independent simulation replicates were used.

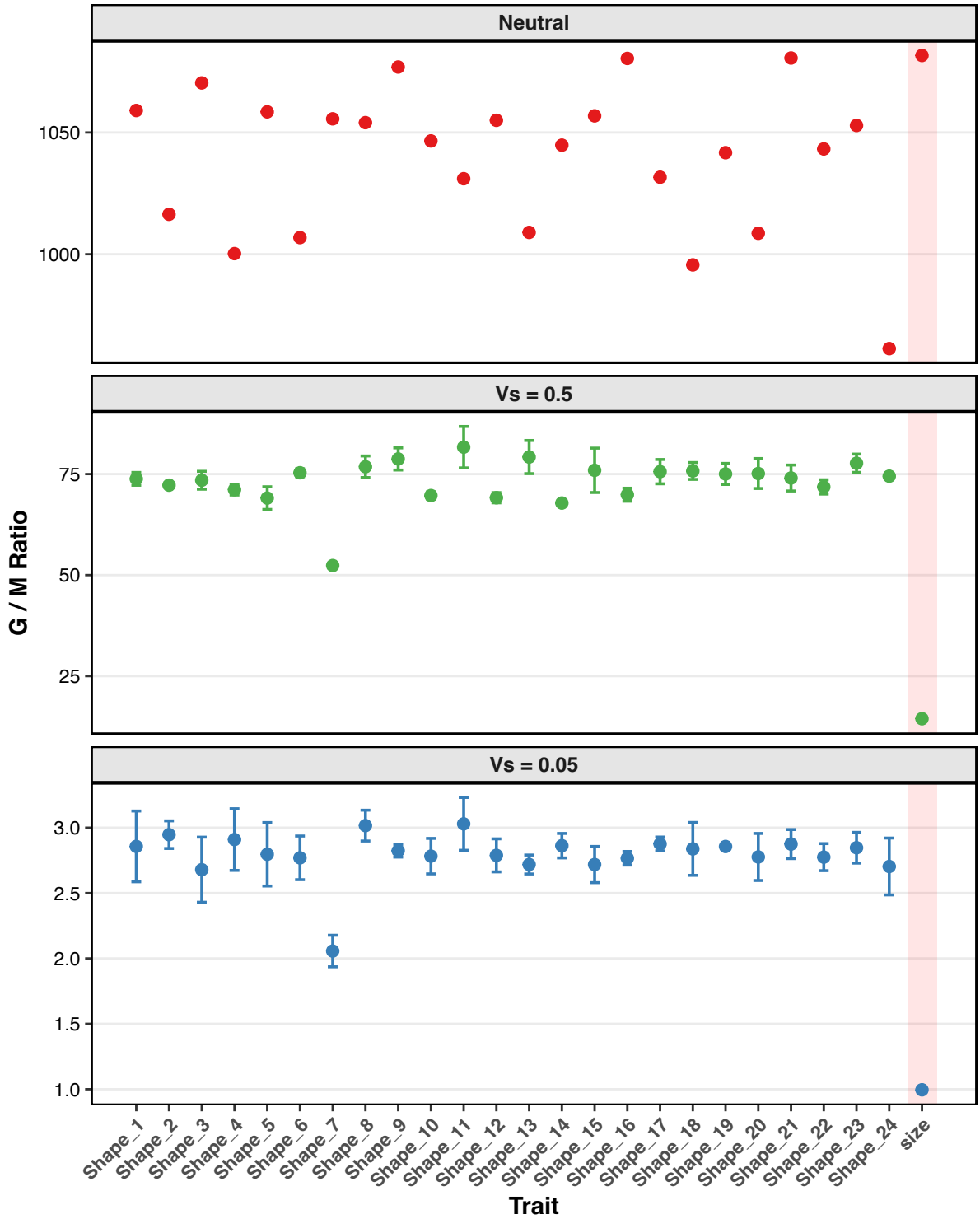

Figure 4: **Ratio of genetic to mutational variance (G/M) across traits under varying evolutionary scenarios.** Points represent the mean G/M ratio across replicates for each trait under Neutral drift, weak stabilizing selection ($V_s = 0.5$), and strong stabilizing selection ($V_s = 0.05$). Error bars indicate ± 1 standard error of the mean across independent simulation replicates ($n = 3$ for selection scenarios). The Neutral panel reflects a single baseline simulation ($n = 1$) and thus lacks error bars. The focal "size" trait is highlighted with a red shaded background to facilitate direct visual comparison of its variance ratio relative to other traits. Note that the y-axis scales vary independently across panels to accommodate differing magnitudes of G/M ratios between scenarios.

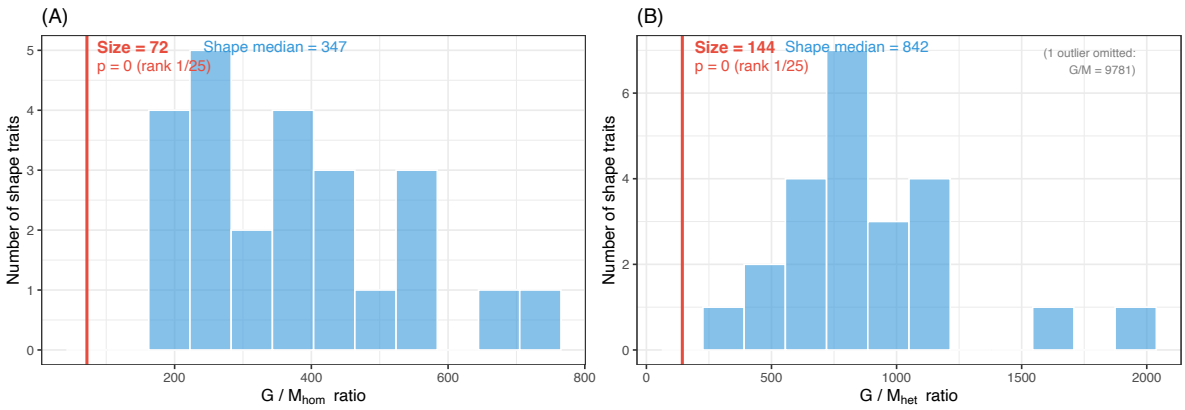

Figure 5: **Empirical G/M ratio for size vs. shape traits in *D. melanogaster*.** Distribution of per-trait G_{ii}/M_{ii} ratios across the 24 shape traits (blue histogram) computed from the empirical G and M matrices of Houle et al. (2017). The vertical red line marks the G/M ratio of centroid size (ln CS). (A) G/M_{hom} (homozygous mutational variance): size ratio = 72, shape median = 347; (B) G/M_{het} (heterozygous mutational variance): size ratio = 144, shape median = 842. Under both estimates, size has the lowest G/M ratio of all 25 traits (rank 1/25), consistent with stronger stabilizing selection on size selectively depleting its standing genetic variance.

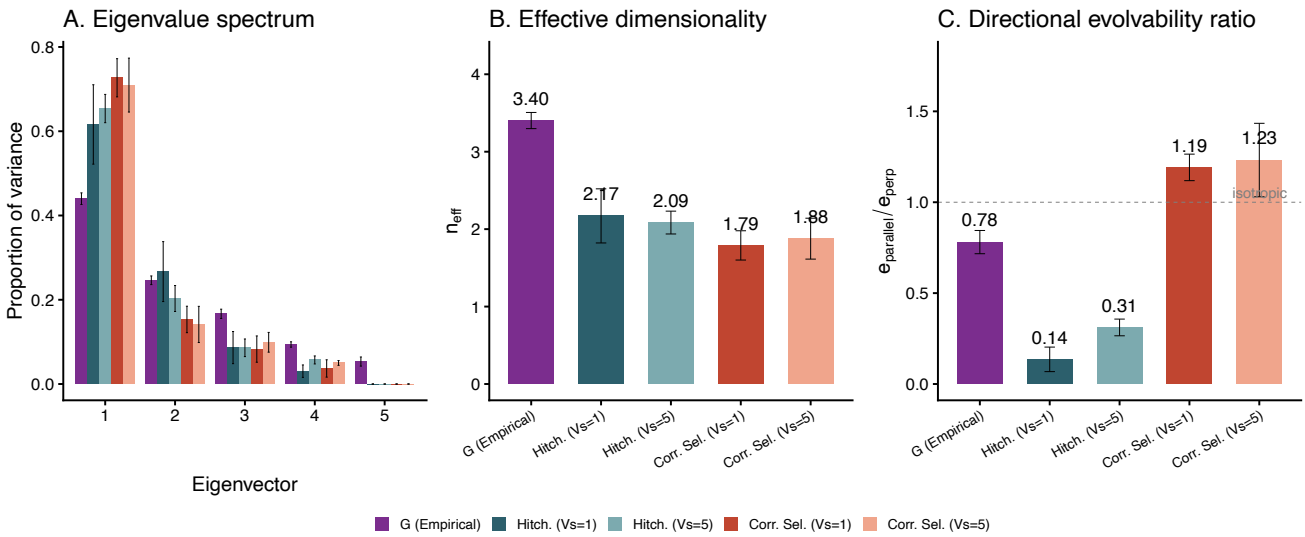

Figure 6: **Eccentricity diagnostics for simulated and empirical G matrices.** Three diagnostics applied to the empirical G from Dugand et al. (2021) and equilibrium G under pleiotropic hitchhiking and correlational selection at two selection strengths ($V_s = 1$ and $V_s = 5$; 5 replicates each, $N = 500$). (A) Eigenvalue spectrum showing the proportion of total genetic variance explained by each eigenvector. (B) Effective dimensionality (n_{eff}), measuring how evenly variance is distributed across eigenvectors. (C) Directional evolvability ratio ($e_{\parallel}/e_{\perp}$), contrasting evolvability along the size direction with the average orthogonal evolvability; the dashed line marks the isotropic baseline (ratio = 1). Error bars for simulated G show ± 1 s.d. across replicates; error bars for empirical G represent estimation uncertainty (parametric bootstrap from REML confidence intervals reported in Dugand et al. 2021).