Identification of multivariate phenotypes most influenced by mutation: *Drosophila serrata* wings as a case study

Cara Conradsen[1,2,3] and Katrina McGuigan[1,4]

1. School of the Environment, The University of Queensland. St. Lucia, QLD, 4072 Australia.
2. European Centre for Environment and Human Health, University of Exeter Medical School, Environment and Sustainability Institute, Treliever Road, Penryn Campus, Penryn, TR10 9FE, United Kingdom.
3. Environment and Sustainability Institute, University of Exeter, Treliever Road, Penryn Campus, Penryn, TR10 9FE, United Kingdom.
4. k.mcguigan1@uq.edu.au

All data analysed in this paper is published: (Conradsen et al., 2021)

Running Title: Multivariate mutational variance

Keywords: pleiotropy mutational variance; mutational covariance; DsGRP

ential.

ABSTRACT

The distribution of pleiotropic mutational effects impacts phenotypic adaptation. However, small effect sizes and high sampling error of covariances hinder investigations of the factors influencing this distribution. Here, we explored the potential for shared information across traits affected by the same mutations to counter sampling error, allowing robust characterisation of patterns of mutational input. Exploiting a published dataset representing 12 samples of the same mutation accumulation experiment in *Drosophila serrata*, we inferred robust signals of mutational effects from the concordance across samples. Krzanowski common subspace analysis identified a multivariate wing trait with statistically supported mutational variance in all samples. Importantly, this multivariate trait was aligned with the major axis of among-line (mutational) variance within most population samples. That is, despite considerable heterogeneity among samples in individual (co)variance parameter estimates, the predominant pattern of correlated mutational effects was identified in datasets reflecting a typical mutation accumulation experimental design. Two other multivariate traits were statistically supported across most samples. Smaller effect sizes (lower mutational variance) with concomitant larger sampling error or other factors (e.g., microenvironmental dependence of effects) may reduce the robustness of estimated mutational input for these traits. Overall, our results suggest sampling error does not preclude multivariate analyses of mutation accumulation experiments from extending our knowledge of pleiotropic mutational effects.

INTRODUCTION

How spontaneous mutations affect the phenotype is central to a range of evolutionary genetic phenomena, including the evolutionary dynamics of genetic variance and the evolutionary fates of populations (Johnson & Barton, 2005; Schultz & Lynch, 1997; Whitlock, 2000; Zhang & Hill, 2002). Reflecting this broad significance, substantial effort has been devoted to characterising the input of new mutation to individual quantitative traits (Conradsen et al., 2022; Halligan & Keightley, 2009; Houle et al., 1996; Walsh & Lynch, 2018). However, this individual trait focus is at odds with the long-standing recognition that selection does not act on individual traits in isolation (Svensson et al., 2021) and that the shared genetic basis of traits influences their response to selection (Walsh & Blows, 2009). Furthermore, theory suggests the strength of correlation of mutational effects across traits is a key determinant of the stability of quantitative genetic covariances, and thus of our ability to predict evolution over multi-generational timescales (Jones et al., 2003).

Several studies have extended beyond individual traits to consider the bivariate distribution of mutational effects, as reflected in the magnitude of trait correlation arising due to novel mutation. Studies estimating mutational correlations between broadly defined reproductive output and survival traits suggest mutational effects on these traits may typically be positively correlated (median mutational correlation = 0.33: Fernandez & Lopez-Fanjul, 1996; Houle et al., 1994; Keightley et al., 2000; Keightley & Ohnishi, 1998; Martorell et al., 1998). Nonetheless, correlations between such traits vary widely in magnitude (-0.5 to 1), and estimates have large confidence intervals. Mutational correlations between morphological, behavioural, locomotor or metabolic traits have also been reported, with estimates spanning moderately negative through strongly positive (e.g., Clark et al., 1995; Estes et al., 2005; Keightley & Ohnishi, 1998; Mackay et al., 1992; Santiago et al., 1992). Where the same traits

have been considered in different studies, as with reproduction and survival traits, marked differences in the mutational correlation estimates are observed (e.g., activity of specific metabolic enzymes: Clark et al., 1995; Harada, 1995). It remains unresolved whether heterogeneity among mutational correlation estimates reflects evolutionarily meaningful differences — such as due to genetic background, taxon or trait type — or instead arises from methodological inconsistencies, differences in trait definitions, or the inherent estimation error associated with quantitative genetic parameters.

As pairwise correlations may offer limited insight into the evolution of phenotypes or genetic variance (Blows & McGuigan, 2015; Johnson & Barton, 2005; Svensson et al., 2021; Walsh & Blows, 2009), it is also important to adopt a broader multivariate (geometric) approach in studies of mutational covariance across larger trait sets. Several more recent studies have interrogated the higher dimensional patterns of correlation in mutational effects captured by the mutational variance covariance matrices (**M**). Intriguingly, some of these studies have found evidence of concordance in major axes between **M** and a corresponding standing genetic variance covariance matrix (**G**) (Dugand et al., 2021; Houle et al., 2017), while others have found marked differentiation between **M** and **G** (Latimer et al., 2014; François Mallard et al., 2023). While further empirical data is essential to clarify the evolutionary implications of these contrasting observations, it is equally critical to improve our understanding of the statistical properties of **M**—especially the sources and consequences of estimation error.

High estimation error is a well-known challenge to studying quantitative genetic parameters (Klein, 1974; Klein et al., 1973; Lynch & Walsh, 1998). Particularly when the number of independent samples (e.g., mutation accumulation lines) is small relative to the number of variables (traits), random sampling error can generate extensive spurious covariance, leading to inflation of the estimated magnitude of variance along leading

eigenvectors of covariance matrices (Blows & McGuigan, 2015; Johnstone, 2007; Sztepanacz & Blows, 2017). The potential influence of sampling error on estimates and conclusions is likely to be exacerbated for mutational relative to standing genetic variation due to the smaller magnitude of (co)variances (i.e., a smaller statistical effect size). For individual traits, mutations contribute, on average, 0.24% to the phenotypic variance per generation (Conradsen et al., 2022; Houle et al., 1996; Lynch et al., 1999), in contrast to the median portion of phenotypic variance due to standing genetic variation, which is >10% across almost any trait type (Hansen et al., 2011). Allowing mutations to accumulate over hundreds of generations of mutation-drift evolution can mitigate the difficulty posed by small magnitude (co)variance (Johnson & Barton, 2005; Lynch et al., 1999). However, estimates from such experiments are increasingly likely to reflect the combined effects of multiple mutations per genotype (mutation accumulation line), which will limit insight into the distribution of effects of individual mutations. Furthermore, logistical limitations might lead to long-term maintenance trading-off with the number of mutationally divergent genotypes (lines) that can be maintained and assayed, opposing the statistical gain from sampling more mutations per genotype.

These concerns about statistical limits may disincentivise investment in the further empirical estimates of **M** that are required to advance our understanding of this important evolutionary parameter. However, by leveraging shared signal among covarying traits (Schmitz et al. 1998), multivariate analyses may, in fact, offer a powerful solution for advancing our understanding of mutational effect distributions. This strategy has proven effective in human genome-wide association studies, where multi-trait frameworks have significantly improved variant detection across diverse traits (Baselmans et al., 2019; Qi & Chatterjee, 2018; Turley et al., 2018; Wu, 2020). Here, we seek to determine whether shared mutational signal might similarly lead to relatively robust identification of the major axis of mutational variation (i.e., the first eigenvector of **M**, $\mathbf{m}_{max}$).

110 To address this question, we take advantage of a unique published dataset in which it
111 is possible to repeatedly estimate **M** for the same mutation accumulation experiment. This data
112 represents sequential samples of *Drosophila serrata* wing phenotypes in 42 lines that have
113 diverged through mutation-drift over 20 generations (Conradsen et al., 2021). There is clear
114 evidence from developmental studies (reviewed in Blair, 2007; Matamoro-Vidal et al., 2015),
115 identification of causal loci (e.g., Pitchers et al., 2019) and estimates of both **G** (McGuigan &
116 Blows, 2007; Mezey & Houle, 2005) and **M** (Dugand et al., 2021; Houle et al., 2017) of
117 pleiotropic effects of alleles across Drosophila wing traits. Thus, we expect traits to share
118 mutational information, essentially resulting in higher replication of phenotypic measures and
119 increased precision of estimation. However, it is unclear whether other factors (e.g.,
120 developmental error, measurement error, micro-environmental effects) also generate shared,
121 but uninteresting, information, negating the multivariate gain in power.

122 **METHODS**

123 **Population history and wing data collection**

124 The experimental design was described in detail in Conradsen et al. (2022), where mutational
125 variance of individual traits was investigated, but no consideration was given to mutational
126 covariance among traits. Briefly, a highly inbred *D. serrata* Reference Genome Panel (DsRGP)
127 line (Reddiex et al., 2018), DsGRP-226, was used as a progenitor to establish 200 Mutation
128 Accumulation (MA) lines (Kannan et al., 2023). After 20 generations of mutation accumulation
129 via mutation-drift in brother-sister inbreed lines, 42 MA lines were randomly chosen and each
130 used to establish two new sublines, one under each of two population size treatments (Small
131 and Large: Figure 1a). For each of the 84 sublines (42 MA x 2 treatments), wings were sampled
132 from up to six male flies in each of two replicate rearing vials for six sequential generations
133 (Figure 1a): a total of 5,135 wings were available for analysis, corresponding to 12 (2

134 treatments by 6 generations) repeated samples in which among-line (mutational) variance could
135 be estimated.

136 The population size treatment, and the sequential sampling introduce the opportunity for
137 segregating new mutations and microenvironment to differentially influence among-line
138 variance estimates, including via environment-specific mutational effects. However,
139 Conradsen et al. (2022) partitioned these potential sources of variance, and provided evidence
140 that sampling error alone accounted for the vast majority of the substantial heterogeneity
141 among the 12 estimates of among-line (mutational) variance, with only a small amount of the
142 variance in one trait (wing size) accounted for by micro-environment and mutation-drift-
143 selection processes. Here, we therefore consider these data as repeated measurement of the
144 effects of mutations fixed in lines during the 20 generations of mutation accumulation, and
145 focus our interpretation on the contribution of sampling error to variation among the 12
146 estimates of mutational covariance parameters.

147 Nine wing landmark positions (Figure 1b) were recorded from images, aligned using a
148 generalised Procrustes least squares superimposition, and centroid size (CS, the square root of
149 the sum of the squared distances between each land-mark and their centroid) determined
150 (Rohlf, 2007). Inter-landmark distances (ILDs, in units of centroid size) were calculated from
151 the aligned coordinates. Using the model described below, we attempted to estimate the among-
152 line variance-covariance matrix for the 11 traits previously analysed individually by Conradsen
153 et al. (2022). However, this model failed to converge – a likely consequence of the relatively
154 few lines (42) and the small magnitude of among-line variance accumulated after 20
155 generations of mutation-drift evolution. We therefore report on a subset of those 11 traits, and
156 consider further in the Discussion how our pragmatic, but non-random, trait selection may
157 influence conclusions. Four (ILD1.9, ILD3.8, ILD4.8 and ILD5.6) of the original 11 traits were

excluded because at least one of the 12 (two treatments measured in six generations) among-line variance estimates was zero (see Figure 5 of Conradsen et al., 2022), causing model convergence failure. A further trait (ILD4.5) also impacted model convergence (models containing this trait returning non-positive definite asymptotic matrices precluding confidence interval estimation) and was therefore also excluded. A final subset of five ILDs, along with CS (Figure 1b), were retained. To facilitate multivariate model convergence, data for each trait was standardised (mean = 0, standard deviation = 1) experiment-wide (across all generations and treatments). Mahalanobis distance was used to identify multivariate outliers (critical value of χ^2 = 22.46, df = 6, α = 0.01), which were excluded from all analyses (55 outliers across the 5135 flies, < 1.1%).

Estimating mutational variance-covariance matrices

We estimated the 12 among-line (**M**) covariance matrices for our six traits using restricted-maximum likelihood (REML). The MIXED procedure in SAS v9.4 (SAS Institute Inc., Cary, NC.) was used to fit:

$$y_{ijk} = \mu + Line_i + Vial_{j(i)} + \varepsilon_{ijk} \quad (1)$$

where y_{ijk} was the vector of six trait values for the kth wing (k = 1…6), for the jth vial (j = 1,2), within the ith line (i = 1...42); μ was the vector of global means per trait; the random effects of *Line*, *Vial* and the residual error (ε) estimated the 6 x 6 unstructured covariance matrices corresponding to among-line (mutational), among replicate vial and among individual flies within a vial, respectively.

To determine the number of statistically supported independent axes of among-line variation, a factor-analytic approach (Hine & Blows, 2006) was applied, using likelihood ratio tests (LRT) to compare sequential models constraining among-line variance from zero up to

181 six dimensions. The LRT degrees of freedom were defined by the difference between models
182 in the number of parameters estimated. Factor analytic modelling conservatively accounts for
183 the effect of sampling error on eigenvalues, representing the most robust approach for
184 determining statistical support for the dimensionality of covariance matrices, (Sztepanacz &
185 Blows, 2017).

186 To place confidence intervals about estimates we applied the REML-MVN method
187 developed by Meyer and Houle (Houle & Meyer, 2015; 2013), drawing 10,000 random
188 samples from the distribution $N \sim (\widehat{\boldsymbol{\theta}}, \mathbf{V})$, where $\widehat{\boldsymbol{\theta}}$ was the vector of observed (REML)
189 covariance parameter estimates, and **V** was the asymptotic variance-covariance matrix. These
190 RML-MVN samples were used to calculate confidence intervals for both model-estimated and
191 derived parameters (detailed below).

192 **Constructing null model variance-covariance matrices**

193 The overarching null hypothesis of the current study is that we have 12 repeated measures of
194 the same mutational (co)variance of wing traits, differing from one another only through
195 transient segregation of recent mutations, micro-environmental effects and estimation error.
196 We constructed a null distribution reflecting this by randomly sampling MA lines (with
197 replacement) to create 12 experimental populations (two treatments by six generations) of 42
198 lines. By retaining the within-line information (i.e., the set of ~12 flies in replicate vials
199 sampled from a subline each generation were collectively assigned their new, random,
200 population identity) while shuffling lines among population size treatments and generations,
201 we retained the signal of mutation in the populations, focusing our null on the potential for
202 sampling error to generate spurious divergence among 12 repeated measures of **M**. We
203 generated 1000 (randomised) null datasets, and again implemented model 1 to estimate 12 **M**
204 in each of these datasets. Some of the 12,000 null data models failed to converge, resulting in

nine of the 1000 null datasets (0.9%) with 11, not 12, estimates of **M**. The null matrices were taken through the same analytical pipeline (detailed below) as the observed data.

Comparing variance-covariance matrices

We used Krzanowski's subspace method to determine the extent to which the 12 **M** estimated the same multivariate axes of mutational variance. This approach is described in detail elsewhere (Aguirre et al., 2014; Blows et al., 2004; Krzanowski, 1979). Briefly, R (R Core Team, 2024) was used to diagonalise each **M**, and to estimate the Krzanowski subspace matrix, **H**:

$$\mathbf{H} = \sum_{t=1}^{p} \mathbf{L}_t^{\mathrm{T}} \mathbf{L}_t \qquad (2)$$

where $\mathbf{L}_t$ contains the subset of the jth largest eigenvectors, as rows, of the tth **M** matrix, (and $\mathbf{L}_t^{\mathrm{T}}$ its transpose containing eigenvectors as columns) and $p = 12$ is the number of populations to be compared. Spectral decomposition of **H** describes the degree of shared structure across the matrices, with the eigenvalues ranging from 0 to p. Here, an eigenvalue of 12 would indicate that all 12 **M** have variance in the direction of the associated eigenvector of **H**. The interpretation of **H**'s eigenvalues depends on how j is defined (Aguirre et al., 2014; Blows et al., 2004; Krzanowski, 1979). When $j > \frac{n}{2}$ (where n is the number of traits; here, $n = 6$, setting $j > 3$), the analysis is constrained to recover common subspaces of two matrices. Conversely, if j is allowed to vary among populations, above $j = \min(j_1, j_2, \ldots j_p)$ dimensions are constrained to be orthogonal (i.e., not contain a common subspace) for at least one population — the one with the lowest j. Here, we followed Aguirre et al. (2014) in defining j for the tth **M** matrix as the number of eigenvectors required to explain at least 90% of the variance; for each of our 12 sampled matrices, this included all dimensions with statistically supported variation (Table 1, Table S2). The 90% criterion ensured that we could interpret at least the first two eigenvalues

228 of **H** (Table 1). For four **M**, $j = 2$, indicating that these matrices were constrained to share no
229 more than two common dimensions with any other matrix. Notably, only one **M** had $j > 3$,
230 indicating that no two matrices were constrained to share common subspaces.

231 To place confidence intervals on the eigenvalues of **H**, we applied equation 2, and
232 extracted the eigenvalues of the resulting matrix, to each of the 10,000 REML samples of the
233 12 matrices. To determine which of the 12 matrices had variance in which parts of the subspace
234 captured by **H**, and whether matrices differed in the magnitude of among-line variance in
235 shared (common) subspaces, we projected the eigenvectors of **H** through each **M** (Aguirre et
236 al., 2014). To place confidence intervals on these estimates of common-subspace variance,
237 eigenvectors of the observed REML **H** matrix were projected through each of the 10,000
238 REML-MVN samples of the matrix. Equation 2 was also applied to estimate **H** for each of the
239 1000 null datasets (we used $j = 3$ eigenvectors for all null datasets). The 95% confidence
240 intervals of eigenvalues of the null data **H** were determined directly from distribution across
241 the null 1000 datasets.

242 **RESULTS**

243 **Among-line variation in the 12 population samples**

244 Consistent with previous analyses of each trait individually (Conradsen et al., 2022), we
245 observed the 12 population estimates of among-line (mutational) variance to vary in both
246 magnitude and statistical support for all traits (Figure 2a,b; Table S1). There were no patterns
247 of consistent change over time or divergence between the population size treatments (Figure
248 2a), suggesting the observed heterogeneity was not due to mutation-drift-selection processes.
249 There was statistical support for 46 (63.9%) of the 72 estimates (6 traits in 12 population
250 samples), with statistical support across all 12 samples for only one trait (ILD1.5, capturing
251 proximo-posterior wing width: Figure 1b) (Figure 2a,b). Particularly for traits with fewer

252 statistically supported estimates (Figure 2b) there were large differences in magnitude of
253 among-line variance (ratio of largest to smallest ranged from 1.93 up to 34.4, median 7.6;
254 Figure 2a). Nonetheless, there was little statistical support for among-line variance to differ
255 between the population samples - estimates typically fell well within each other's 90% CI, with
256 only a few exceptions for traits ILD1.2 (two largest versus five smallest estimates) and ILD2.8
257 (largest versus smallest) (Figure 2a).

258 There was considerably less statistical support for mutational covariance among traits,
259 with only 15% (27 of 180) of pairwise covariances statistically distinct from zero (Figure 2a,c).
260 Nonetheless, the average absolute correlation was 0.51 (Table S1). This pattern of frequent
261 strong association between traits suggests that the relatively large sampling error of covariance
262 estimates, rather than a lack of correlated mutational effects among traits, might contribute to
263 the limited statistical evidence for mutational covariance. A binomial test of the null hypothesis
264 of an equal number of positive and negative values rejects the null hypothesis (P = 0.0193)
265 when 10 of 12 estimates have a consistent sign; six (of 15, 40%) mutational correlations were
266 thus statistically supported as non-zero (Figure 2a,c) across the total dataset. There was,
267 however, marked heterogeneity among the 12 repeated estimates of each pairwise covariance
268 (correlation). The correlation estimates spanned the full range (-1 to 1) for one trait pair
269 (ILD1.2: ILD2.8), with non-overlapping 95% CI of estimates for this and two other trait pairs
270 (CS: ILD1.2 and CS: ILD3.7) (Figure 2a,c; Table S1).

271 **Common subspace and major axis of mutational variation**

272 The distribution of eigenvalues of each population sample's among-line covariance matrix was
273 consistent with mutation generating trait covariance, and with variation in mutational effects
274 among traits. Specifically, the large (relative to variance of individual traits) first eigenvalue
275 (Figure 3; Table 1 vs Table S1) suggests traits share variation, while the robust statistical

support for two axes of mutational variance in most population samples (Table S2) suggested that not all mutations generated the same correlated effects across traits. Here, we consider whether the similar pattern of eigenvalues (Figure 3) and low (relative to individual traits) heterogeneity among the 12 replicate samples in their eigenvalues (ratio of largest to smallest population sample estimate: 1.92, 2.90 and 3.08 for m_1, m_2 and m_3, respectively: Table 1) also reflect concordance of the associated eigenvectors.

The eigenvalue of the major axis (h_1) of the Krzanowski common subspace was within the confidence bounds estimated from the null datasets simulating a single population, and close to the maximum value (12) observed when the vector is present in the sampled subspaces of all 12 populations (Table 2). Consistent with this interpretation that each of the 12 populations had estimated the same multivariate axis of mutational input, there was statistically supported variance along h_1 in all 12 population samples (Figure 4a). This common axis of wing shape variation, representing the best estimate of the multivariate wing shape most affected by mutation in the experiment, reflected an axis of expansion versus contraction of the proximal versus distal wing length (strongest, opposing, loadings on h_1 from ILD1.5 and ILD3.7: Table 2). Notably, these two traits had the most consistent statistical support for, and least heterogeneity in magnitude among, among-line variance estimates, and were significantly negatively correlated in all but one population sample (average bounded correlation coefficient = -0.87: Figure 2; Table S1).

The major axis of multivariate mutational input, h_1, was typically captured by the leading eigenvector of a population sample's **M** (i.e., was aligned with m_1). This is evidenced by: i) h_1 being most closely aligned to m_1 in 10 of the 12 populations (Table S4); ii) the magnitude of among-line variance being greater along h_1 than along h_2 or h_3 in 11 populations (Figure 4a,b) and; iii) h_1 typically accounting for almost as much variance in a population's

300 covariance matrix as m_1 did (median = 86% of m_1: Figure 4d). This reduced variance along h_1
301 relative to each m_1 is expected as variance along m_1 will be inflated by estimation error (Blows
302 & McGuigan, 2015), unique to that specific population sample. In population S2, h_1 was more
303 closely aligned with m_2 (Table S4), reflected in less variance along h_1 than h_2 (0.30 versus 0.36:
304 Figure 4a,b) and h_1 accounting for only 60% as much variance as m_1 (lower right point in
305 Figure 4d). Nonetheless, even for this relatively incongruent population sample, m_1 did capture
306 a substantial signal of the major axis of mutational variance (h_1) (Figure 4a,d). Similar to the
307 observation of reduced heterogeneity (relative to individual traits) among eigenvalues of **M**,
308 the largest estimate of among-line variance along h_1 was 1.96 times greater than the smallest
309 estimate, with estimates typically well within the CI of all other population samples (Figure
310 4a).

311 The eigenvalues (8.9 and 6.6: Table 2) of the other two interpretable common subspace
312 axes implicated further axes of mutational input, with non-zero variance along these axes in
313 most populations (Figure 4b,c). However, the absence of statistically supported variance along
314 these axes in some population samples suggests that robust estimation of this variation may be
315 limited by factors such as estimation error (increased relative to h_1 due to the smaller magnitude
316 of effect), segregating mutations, or micro-environmental heterogeneity. There were no clear
317 temporal or among population size treatment patterns in the magnitude of variance along h_2 or
318 h_3 (Figure 4b,c; Table S4). Notably, h_2 primarily reflected mutational variance in wing size
319 (CS), while h_3 was most strongly influenced by ILD2.5 (Table 2): these are the only two traits
320 for which individual trait analyses presented by Conradsen et al. (2022) robustly rejected the
321 null hypothesis that among-line variance was homogenous among generations or population
322 size treatments. ILD1.2 and ILD2.8 also contributed to h_2 (larger versus smaller wings had
323 disproportionally longer versus shorter anterior margins: Table 2, Figure 1b). Low concordance
324 among the 12 estimates of variances and covariance involving these traits (Figure 2) suggest

325 we have least confidence in the mutational effects on these traits, but shared signal (i.e.,
326 correlated pleiotropic effects) with other traits may underpin their contribution to this axis of
327 wing shape variation, observed in most population samples (Figure 4b).

328 **DISCUSSION**

329 Accurately estimating mutational effects, and particularly pleiotropic effects across traits,
330 represents a clear challenge to understanding how factors such as genetic background,
331 environmental context or the traits under consideration might each influence mutational
332 characteristics (Conradsen et al., 2022). However, by leveraging shared signal against
333 independent error, multivariate analyses of mutational covariance across traits may potentially
334 provide more robust insights than interrogation of mutational correlations of individual trait
335 pairs. Here, we provide support for this through re-analysis of published data from an
336 experiment that was: i) designed to reflect typical levels of sampling error of a spontaneous
337 mutation accumulation experiment (Conradsen et al., 2022) and; ii) is in a system (Drosophild
338 wings) where shared mutational signal is expected (extensive developmental and genetic
339 evidence of pleiotropic alleles and trait covariances: Carreira et al., 2011; Dugand et al., 2021;
340 Houle et al., 2017; Matamoro-Vidal et al., 2015; Pitchers et al., 2019). Our results suggest that
341 while mutational differences may be detectable for major axes of mutational variance, for other
342 multivariate traits this may remain difficult. We discuss these points below, and highlight the
343 value of replicated estimation for deepening our understanding of mutational characteristics.

344 Estimating mutational (co)variances repeatedly from independent samples of the same
345 mutation accumulation experiment allowed us to confidently conclude that there was a strong
346 signal of mutation generating trait covariance. Notably, although all 12 matrices captured
347 variation along the same axis of wing shape (h_1: Table 2; Figure 4a), the trait covariance
348 contributing most strongly to this robustly estimated axis of mutational variance was not

349 unequivocally identifiable *a priori* from a single population sample. We expected the
350 magnitude of mutational (co)variance would predict the extent of concordance across the 12
351 estimates. However, the covariance contributing most strongly to h_1 (ILD1.5: ILD3.7), while
352 the most consistently statistically supported across samples (Figure 2c), was the strongest of
353 the statistically supported trait correlations in only one population sample. Similarly, while the
354 two traits individually had relatively large average estimates of among-line variance, other
355 traits had the largest (statistically supported) estimate in some samples (Figure 2; Tables S1)
356 (also see Figure 7 versus Figure 5 of Conradsen et al., 2022). Thus, the magnitude of estimated
357 (co)variance in any single sample of mutational phenotypes may not provide a simple
358 diagnostic of the strongest mutational signal.

359 Among-line variation along h_1 was predominantly captured by variation along the
360 major axis of the among-line covariance matrix estimated within each population sample (i.e.,
361 m_1). Higher variance along population-specific m_1 than along h_1 (Figure 4d) is consistent with
362 the Krzanowski common subspace analysis finding the shared signal of mutational input, with
363 independent sampling error within each population sample leading to overestimation of the
364 magnitude of mutational variance associated with the most mutationally variable multivariate
365 trait. Leading eigenvalues of genetic (and mutational) covariance matrices are particularly
366 vulnerable to inflation by sampling error (Blows & McGuigan, 2015; Hill & Thompson, 1978;
367 McGuigan et al., 2014). Statistical approaches can estimate the inflation on eigenvalues
368 (McGuigan et al., 2014), although the empirical approach of repeated estimation from
369 independent samples applied here will provide greater certainty as to the orientation of the
370 eigenvectors.

371 Here, concordant patterns across replicate independent samples of the same mutation
372 accumulation experiment provided considerable confidence that these patterns reflect the

373 underlying mutational signal. We suggest that concordance of mutational covariances (or major
374 axes of mutational variance) across independent samples of the mutational spectrum (e.g.,
375 independent mutation accumulation experiments) may help distinguish the contribution of
376 pleiotropy from that of linkage, leading to deeper insight into general characteristics of the
377 distribution of pleiotropic effects. In mutation accumulation experiments, all mutations arising
378 within a line are in linkage. Particularly for traits with directionally biased mutational effects
379 (e.g., fitness traits, which evolve lower values under mutation-drift: Halligan & Keightley,
380 2009), linkage of mutations with independent, but directionally biased, effects on each trait
381 may contribute substantially to among-line covariances (Estes et al., 2005). For traits without
382 biased directional effects, including *Drosophila* wing traits (Houle & Fierst, 2013; Santiago et
383 al., 1992), strong trait covariances may still arise via linkage of randomly sampled mutations.
384 However, chance linkage of randomly sampled mutations is not expected to cause consistent
385 covariance patterns across independent mutation accumulation experiments. Rather, in the
386 absence of directional bias to the distribution of mutational effects, the average mutational
387 covariance generated by linkage should approach zero. Repeated observation of the same major
388 axes of phenotypic variation in independent samples of mutation therefore suggest conserved
389 patterns of correlated pleiotropic effects of mutations.

390 The major axes of mutational input to wing shape estimated in the current study (h_1,
391 Table 2) suggests that mutation predominantly generated variation along an axis of shorter and
392 wider versus longer and narrower wing tips. Intriguingly, this was qualitatively the same
393 mutational wing shape variance identified in independent studies of *Drosophila* wings, using
394 different experimental designs and analytical approaches. Dugand et al. (2021) estimated
395 mutational (co)variance in an outbred, randomly mating, *D. serrata* population founded by the
396 DsGRP lines (including the ancestor line of the mutation accumulation lines considered in the
397 current study). In that outbred population analysis, m_1 was characterized by lengthening ILD1.5

398 and ILD5.7 while shortening along ILD4.6 and ILD3.9 (and vice versa: see their Table 3).
399 Similarly, Houle et al. (2017) estimated the major axis of mutational variance using inbred
400 mutation accumulation lines of *D. melanogaster* (see their Extended Data Figure 1), finding
401 that mutation pushed out the landmarks corresponding to LM5 and LM7 in the current study
402 (increasing ILD1.5 and ILD1.7), while pulling closer together landmarks corresponding to
403 LM3,4 and LM8,9 (shortening distances between those landmarks, including ILD3.7, ILD3.9
404 and ILD4.6).

405 The apparent consistency of multivariate mutational covariance patterns across these
406 different studies of *Drosophila* wings suggests bias in either the frequency or effect sizes of
407 pleiotropic mutations, driving persistent bias in the phenotypic variation generated by new
408 mutation. This inference of conserved mutational effects is qualitative rather than quantitative.
409 Conservation of mutational effects among taxa is also implicated by evidence that the
410 mutational covariance matrix estimated in one species (*D. melanogaster*) strongly predicted
411 the pattern of divergence among 112 Drosophilidae species over 40 million years (Houle et al.,
412 2020; Houle et al., 2017). In contrast, Houle and Fierst (2013) compared **M** estimated from
413 two genotypes of *D. melanogaster* and concluded that these **M** were quite different from one
414 another, including in the orientation of m_1. Few other studies have quantitatively compared **M**
415 estimated from independent samples of mutations. François Mallard et al. (2023) estimated **M**
416 of locomotor behavior traits in two genotypes of *Caenorhabditis elegans*, finding no evidence
417 that these independently sampled **M** diverged.

418 Direct, quantitative, comparison of **M** from different studies is hampered by
419 methodological differences (here, for example, differences in wing landmarks visible under
420 different imaging methods). Inconsistent statistical support for mutational variance of
421 individual traits can also limit direct comparability. Although Conradsen et al. (2022) measured

422 the same traits as Dugand et al. (2021), differences in which traits had non-zero mutational
423 variance estimates, coupled with the impact of zero variance on multivariate model
424 convergence, led to different traits being included in analyses of **M**. Whether inconsistency
425 among independent studies in the detection of mutational variance for individual traits reveals
426 genotypic differences in mutational effects will require further comparative data to assess the
427 role of sampling error.

428 Even when mutational variance is observed for individual traits, mutational variance
429 co-variance matrices are characterised by so-called nearly null spaces (Gomulkiewicz & Houle,
430 2009) – that is, **M** typically have one or more very small eigenvalues, statistically
431 indistinguishable from zero (Dugand et al., 2021; Estes et al., 2005; Hine et al., 2018; Houle &
432 Fierst, 2013; Latimer et al., 2014; François Mallard et al., 2023; Miller et al., 2023). While this
433 evidence that mutational input to phenotypic variation is uneven across phenotypic trait space
434 is interesting on its own, it is challenging to investigate further the traits with low mutational
435 input. Comparative approaches may again extend our understanding of this question. Houle
436 and Fierst (2013) found that, despite differences between *D. melanogaster* genotypes in the
437 multivariate traits with high mutational input, the traits associated with low mutation were
438 concordant. Dugand et al. (2021) found that *D. serrata* wing phenotypes with low
439 (undetectable) mutational input also have very low standing genetic variance. These
440 observations suggest that it is not simply chance sampling of mutation or estimation error that
441 causes low mutational input to some multivariate traits, and opens the path to further investigate
442 why such variation might be relatively inaccessible.

443 Overall, our results suggest that even relatively modest mutation accumulation
444 experiments may provide valuable information on the multivariate axes of traits space to which
445 mutation contributes the greatest new variance each generation. Such insights are valuable,

446 particularly for deepening our understanding of how mutation and selection interact to shape
447 standing genetic variance, and long-term responses to selection – for example, does mutation
448 predominantly introduce variance to traits with relatively high standing genetic variation (as
449 observed for Drosophildae wings: Dugand et al., 2021; Houle et al., 2017) or is there a
450 mismatch between mutational and standing genetic variance implicating selection acting
451 against new mutations (for example, as observed for thermal performance traits: Latimer et al.,
452 2014; or locomotor behaviour traits: F Mallard et al., 2023). Replication of estimates both
453 within the same experiment and across mutation accumulation experiments are a powerful
454 approach to both identify dominant patterns of pleiotropic correlations, and to ensure robust
455 inference of how that distribution of effects may vary.

456

DATA AVAILABILITY

Data analysed in this paper were published by (Conradsen et al., 2021)

ACKNOWLEDGMENTS

The authors thank Stephen F. Chenoweth for providing us with the *D. serrata* MA lines, and Adam Reddiex, Nicholas Appleton, Jack Price and Derek Sun for their help with maintaining the flies and collecting the data. The authors also thank Emma Hine for helpful discussions about analyses.

CONFLICTS OF INTEREST

None declared.

469 **LITERATURE CITED**

470 Aguirre, J. D., Hine, E., McGuigan, K., & Blows, M. W. (2014). Comparing G: multivariate analysis of genetic
471 variation in multiple populations. *Heredity*, *112*(1), 21-29. https://doi.org/10.1038/hdy.2013.12

472 Baselmans, B. M., Jansen, R., Ip, H. F., van Dongen, J., Abdellaoui, A., van de Weijer, M. P., Bao, Y., Smart, M.,
473 Kumari, M., & Willemsen, G. (2019). Multivariate genome-wide analyses of the well-being spectrum.
474 *Nature Genetics*, *51*(3), 445-451.

475 Blair, S. S. (2007). Wing vein Patterning in *Drosophila* and the analysis of intercellular signaling. *Annual Review*
476 *of Cell and Developmental Biology*, *23*, 293-319.
477 https://doi.org/10.1146/annurev.cellbio.23.090506.123606

478 Blows, M. W., Chenoweth, S. F., & Hine, E. (2004). Orientation of the genetic variance-covariance matrix and the
479 fitness surface for multiple male sexually selected traits. *Am Nat*, *163*(3), 329-340.
480 https://doi.org/10.1086/381941

481 Blows, M. W., & McGuigan, K. (2015). The distribution of genetic variance across phenotypic space and the
482 response to selection. *Molecular Ecology*, *24*(9), 2056-2072. https://doi.org/10.1111/mec.13023

483 Carreira, V. P., Soto, I. M., Mensch, J., & Fanara, J. J. (2011). Genetic basis of wing morphogenesis in *Drosophila*:
484 Sexual dimorphism and non-allometric effects of shape variation. *BMC Developmental Biology*, *11*, 32.
485 https://doi.org/10.1186/1471-213x-11-32

486 Clark, A. G., Wang, L., & Hulleberg, T. (1995). Spontaneous mutation-rate of modifiers of metabolism in
487 *Drosophila*. *Genetics*, *139*(2), 767-779. <Go to ISI>://A1995QE14500020

488 Conradsen, C., Blows, M. W., & McGuigan, K. (2021). Variability of Mutational Variance Estimates:
489 MA_dserrata_wings. *Figshare*, *https://doi.org/10.6084/m9.figshare.14913051.v3*. https://doi.org/
490 https://doi.org/10.6084/m9.figshare.14913051.v3

491 Conradsen, C., Blows, M. W., & McGuigan, K. (2022). Causes of variability in estimates of mutational variance
492 from mutation accumulation experiments. *Genetics*, *221*(2), iyac060.
493 https://doi.org/10.1093/genetics/iyac060

494 Dugand, R. J., Aguirre, J. D., Hine, E., Blows, M. W., & McGuigan, K. (2021). The contribution of mutation and
495 selection to multivariate quantitative genetic variance in an outbred population of *Drosophila serrata*.
496 *Proceedings of the National Academy of Sciences*, *118*(31), e2026217118.

497 Estes, S., Ajie, B. C., Lynch, M., & Phillips, P. C. (2005). Spontaneous mutational correlations for life-history,
498 morphological and behavioral characters in *Caenorhabditis elegans*. *Genetics*, *170*(2), 645-653.

499 Fernandez, J., & Lopez-Fanjul, C. (1996). Spontaneous mutational variances and covariances for fitness-related
500 traits in *Drosophila melanogaster*. *Genetics*, *143*(2), 829-837. <Go to ISI>://A1996UN68400019

501 Gomulkiewicz, R., & Houle, D. (2009). Demographic and genetic constraints on evolution. *The American*
502 *Naturalist*, *174*(6), E218-E229. https://doi.org/10.1086/645086

503 Halligan, D. L., & Keightley, P. D. (2009). Spontaneous mutation accumulation studies in evolutionary genetics.
504 *Annual Review of Ecology, Evolution and Systematics*, *40*, 151-172. https://doi.org/DOI
505 10.1146/annurev.ecolsys.39.110707.173437

506 Hansen, T. F., Pelabon, C., & Houle, D. (2011). Heritability is not evolvability. *Evolutionary Biology*, *38*(3), 258-
507 277. https://doi.org/10.1007/s11692-011-9127-6

508 Harada, K. (1995). A quantitative analysis of modifier mutations which occur in mutation accumulation lines in
509 *Drosophila melanogaster*. *Heredity*, *75*, 589-598. https://doi.org/DOI10.1038/hdy.1995.178

510 Hill, W. G., & Thompson, R. (1978). Probabilities of non-positive definite between-group or genetic covariance
511 matrices. *Biometrics*, *34*, 429-439.

512 Hine, E., & Blows, M. W. (2006). Determining the effective dimensionality of the genetic variance–covariance
513 matrix. *Genetics*, *173*(2), 1135-1144.

Hine, E., Runcie, D. E., McGuigan, K., & Blows, M. W. (2018). Uneven distribution of mutational variance across the transcriptome of *Drosophila serrata* revealed by high-dimensional analysis of gene expression. *Genetics*, *209*(4), 1319-1328. https://doi.org/10.1534/genetics.118.300757

Houle, D., Bolstad, G. H., & Hansen, T. F. (2020). Fly wing evolutionary rate is a near-isometric function of mutational variation. *bioRxiv*, 2020.2008.2027.268938. https://doi.org/10.1101/2020.08.27.268938

Houle, D., Bolstad, G. H., van der Linde, K., & Hansen, T. F. (2017). Mutation predicts 40 million years of fly wing evolution. *Nature*, *548*(7668), 447-450.

Houle, D., & Fierst, J. (2013). Properties of spontaneous mutational variance and covariance for wing size and shape in *Drosophila melanogaster*. *Evolution*, *67*(4), 1116-1130. https://doi.org/10.1111/j.1558-5646.2012.01838.x

Houle, D., Hughes, K. A., Hoffmaster, D. K., Ihara, J., Assimacopoulos, S., & Charlesworth, B. (1994). The effects of spontaneous mutation on quantitative traits. I. Variances and covariances of life history traits. *Genetics*, *138*(3), 773-785.

Houle, D., & Meyer, K. (2015). Estimating sampling error of evolutionary statistics based on genetic covariance matrices using maximum likelihood. *Journal of Evolutionary Biology*, *28*(8), 1542-1549.

Houle, D., Morikawa, B., & Lynch, M. (1996). Comparing mutational variabilities. *Genetics*, *143*(3), 1467-1483.

Johnson, T., & Barton, N. (2005). Theoretical models of selection and mutation on quantitative traits. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1459), 1411-1425. https://doi.org/DOI 10.1098/rstb.2005.1667

Johnstone, I. M. (2007). High dimensional statistical inference and random matrices. *Proceedings of the International Congress of Mathematics*, 307-333.

Jones, A. G., Arnold, S. J., Bürger, R., & Houle, D. (2003). Stability of the G-matrix in a population experiencing pleiotropic mutation, stabilizing selection, and genetic drift. *Evolution*, *57*(8), 1747-1760.

Kannan, A., Dugand, R. J., Appleton, N. C., Chenoweth, S. F., Sgro, C. M., & McGuigan, K. (2023). Environmental dependence of mutational (co)variances of adaptive traits. *Evolution*, *77*(11), 2341–2351. https://doi.org/10.1093/evolut/qpad154

Keightley, P. D., Davies, E. K., Peters, A. D., & Shaw, R. G. (2000). Properties of ethylmethane sulfonate-induced mutations affecting life-history traits in *Caenorhabditis elegans* and inferences about bivariate distributions of mutation effects. *Genetics*, *156*(1), 143-154.

Keightley, P. D., & Ohnishi, O. (1998). EMS-induced polygenic mutation rates for nine quantitative characters in *Drosophila melanogaster*. *Genetics*, *148*(2), 753-766. <Go to ISI>://WOS:000072187500019

Klein, T. W. (1974). Heritability and genetic correlation: statistical power, population comparisons, and sample size. *Behavior Genetics*, *4*(2), 171-189.

Klein, T. W., DeFries, J., & Finkbeiner, C. T. (1973). Heritability and genetic correlation: standard errors of estimates and sample size. *Behavior Genetics*, *3*(4), 355-364.

Krzanowski, W. (1979). Between-groups comparison of principal components. *Journal of the American Statistical Association*, *74*(367), 703-707.

Latimer, C. A. L., McGuigan, K., Wilson, R. S., Blows, M. W., & Chenoweth, S. F. (2014). The contribution of spontaneous mutations to thermal sensitivity curve variation in *Drosophila serrata*. *Evolution*, *68*(6), 1824-1837. https://doi.org/Doi 10.1111/Evo.12392

Lynch, M., Blanchard, J., Houle, D., Kibota, T., Schultz, S., Vassilieva, L., & Willis, J. (1999). Spontaneous deleterious mutation. *Evolution*, *53*(3), 645-663. <Go to ISI>://000081507300001

Lynch, M., & Walsh, B. (1998). Correlations between characters. In *Genetics and Analysis of Quantitative Traits* (pp. 629-656). Sinauer.

Mackay, T. F. C., Lyman, R. F., & Jackson, M. S. (1992). Effects of P-element insertions on quantitative traits in *Drosophila melanogaster*. *Genetics*, *130*(2), 315-332. <Go to ISI>://WOS:A1992HB83000009

Mallard, F., Afonso, B., & Teotónio, H. (2023). Selection and the direction of phenotypic evolution. *Elife*, *12*, Article e80993. https://doi.org/10.7554/eLife.80993

Mallard, F., Noble, L., Baer, C. F., & Teotónio, H. (2023). Variation in mutational (co) variances. *G3: Genes, Genomes, Genetics*, *13*(2), jkac335.

Martorell, C., Toro, M. A., & Gallego, C. (1998). Spontaneous mutation for life-history traits in Drosophila melanogaster. *Genetica*, *102-3*, 315-324. <Go to ISI>://WOS:000075404900026

Matamoro-Vidal, A., Salazar-Ciudad, I., & Houle, D. (2015). Making quantitative morphological variation from basic developmental processes: Where are we? The case of the *Drosophila* wing. *Developmental Dynamics*, *244*(9), 1058-1073. https://doi.org/10.1002/dvdy.24255

McGuigan, K., & Blows, M. W. (2007). The phenotypic and genetic covariance structure of Drosophilid wings. *Evolution*, *61*(4), 902-911.

McGuigan, K., Collet, J. M., Allen, S. L., Chenoweth, S. F., & Blows, M. W. (2014). Pleiotropic mutations are subject to strong stabilizing selection. *Genetics*, *197*(3), 1051-1062.

Meyer, K., & Houle, D. (2013). Sampling based approximation of confidence intervals for functions of genetic covariance matrices. *Proceedings of the Association for the Advancement of Animal Breeding and Genetics*, *20*, 523-526.

Mezey, J. G., & Houle, D. (2005). The dimensionality of genetic variation for wing shape in *Drosophila melanogaster Evolution*, *59*(5), 1027-1038. https://doi.org/10.1554/04-491

Miller, C. L., Sun, D. R., Thornton, L. H., & McGuigan, K. (2023). The contribution of mutation to variation in temperature-dependent sprint speed in zebrafish, *Danio rerio*. *American Naturalist*, *202*(4), 519-533. https://doi.org/10.1086/726011

Pitchers, W., Nye, J., Márquez, E. J., Kowalski, A., Dworkin, I., & Houle, D. (2019). A multivariate Genome-Wide Association Study of wing shape in *Drosophila melanogaster*. *Genetics*, *211*(4), 1429-1447. https://doi.org/10.1534/genetics.118.301342

Qi, G., & Chatterjee, N. (2018). Heritability informed power optimization (HIPO) leads to enhanced detection of genetic associations across multiple traits. *PLoS Genetics*, *14*(10), e1007549.

R Core Team. (2024). R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria.*, *https://www.R-project.org/*.

Reddiex, A. J., Allen, S. L., & Chenoweth, S. F. (2018). A genomic reference panel for *Drosophila serrata*. *G3: Genes, Genomes, Genetics*, *8*(4), 1335-1346.

Rohlf, F. (2007). tpsRelw version 1.45. *Department of Ecology and Evolution, State University of New York, Stony Brook*.

Santiago, E., Albornoz, J., Dominguez, A., Toro, M. A., & Lopez-Fanjul, C. (1992). The distribution of spontaneous mutations on quantitative traits and fitness in *Drosophila melanogaster*. *Genetics*, *132*(3), 771-781. <Go to ISI>://WOS:A1992JV14400012

Schultz, S. T., & Lynch, M. (1997). Mutation and extinction: The role of variable mutational effects, synergistic epistasis, beneficial mutations, and degree of outcrossing. *Evolution*, *51*(5), 1363-1371. <Go to ISI>://A1997YE98000003

Svensson, E. I., Arnold, S. J., Burger, R., Csillery, K., Draghi, J., Henshaw, J. M., Jones, A. G., De Lisle, S., Marques, D. A., McGuigan, K., Simon, M. N., & Runemark, A. (2021). Correlational selection in the age of genomics. *Nature Ecology & Evolution*, *5*(5), 562-573. https://doi.org/10.1038/s41559-021-01413-3

Sztepanacz, J. L., & Blows, M. W. (2017). Accounting for sampling error in genetic eigenvalues using random matrix theory. *Genetics*, *206*(3), 1271-1284.

Turley, P., Walters, R. K., Maghzian, O., Okbay, A., Lee, J. J., Fontana, M. A., Nguyen-Viet, T. A., Wedow, R., Zacher, M., & Furlotte, N. A. (2018). Multi-trait analysis of genome-wide association summary statistics using MTAG. *Nature Genetics*, *50*(2), 229-237.

Walsh, B., & Blows, M. W. (2009). Abundant genetic variation + strong selection = multivariate genetic constraints: A geometric view of adaptation. *Annual Review of Ecology Evolution and Systematics*, *40*, 41-59. https://doi.org/10.1146/annurev.ecolsys.110308.120232

Walsh, B., & Lynch, M. (2018). Maintenance of quantitative genetic variation. In *Evolution and Selection of Quantitative Traits* (pp. 1015–1078). Oxford University Press.

Whitlock, M. C. (2000). Fixation of new alleles and the extinction of small populations: Drift load, beneficial alleles, and sexual selection. *Evolution*, *54*(6), 1855-1861. <Go to ISI>://000166682500002

Wu, C. (2020). Multi-trait genome-wide analyses of the brain imaging phenotypes in UK Biobank. *Genetics*, *215*(4), 947-958.

Zhang, X. S., & Hill, W. G. (2002). Joint effects of pleiotropic selection and stabilizing selection on the maintenance of quantitative genetic variation at mutation-selection balance. *Genetics*, *162*(1), 459-471. <Go to ISI>://000178363400037

TABLES AND FIGURES

Table 1. The variance associated with independent axes of trait mutational variance. Eigenanalyses were performed separately for **M** in each population size treatment (Treat: Small or Large) in each generation (Gen: 1 through 6). The eigenvalues (i.e., λ) and their 90% CIs* are reported, along with the number of eigenvectors (*j*) accounting for at least 90% of the variation. These *j* eigenvectors were included in the estimation of the Krzanowski subspace. The estimated **M** and the eigenvector loadings are reported in Tables S1 and S3, respectively.

Treat	Gen	m_1	m_2	m_3	m_4	m_5	m_6	*j*
	1	0.495	0.102	0.084	0.027	0.001	-0.024	3
		0.193; 0.798	0.014; 0.191	0.004; 0.165	-0.036; 0.092	-0.025; 0.027	-0.073; 0.026	
	2	0.506	0.270	0.066	0.015	-0.014	-0.050	2
		0.206; 0.804	0.133; 0.402	-0.005; 0.136	-0.013; 0.043	-0.040; 0.012	-0.091; -0.009	
	3	0.276	0.116	0.037	0.016	-0.011	-0.056	2
		0.079; 0.472	0.032; 0.197	-0.028; 0.102	-0.039; 0.071	-0.042; 0.020	-0.138; 0.028	
Small	4	0.279	0.205	0.051	0.019	0.000	-0.013	3
		0.087; 0.471	0.029; 0.375	-0.021; 0.121	-0.008; 0.045	-0.025; 0.025	-0.044; 0.017	
	5	0.429	0.159	0.114	0.029	0.010	-0.07	3
		0.213; 0.648	0.023; 0.297	0.008; 0.219	-0.001; 0.060	-0.015; 0.034	-0.147; 0.007	
	6	0.439	0.093	0.047	0.021	-0.004	-0.037	2
		0.238; 0.641	0.020; 0.165	-0.001; 0.094	-0.006; 0.049	-0.044; 0.034	-0.099; 0.025	
	1	0.446	0.267	0.070	0.033	0.008	-0.007	3
		0.211; 0.676	0.131; 0.406	0.007; 0.136	-0.017; 0.082	-0.013; 0.030	-0.040; 0.026	
	2	0.313	0.215	0.077	0.052	0.025	0.000	4
		0.164; 0.460	0.073; 0.354	0.015; 0.138	-0.006; 0.110	-0.000; 0.050	-0.020; 0.019	
	3	0.529	0.144	0.070	0.020	-0.020	-0.038	2
		0.272; 0.792	-0.014; 0.300	0.004; 0.137	-0.009; 0.051	-0.069; 0.029	-0.115; 0.039	
Large	4	0.303	0.160	0.101	0.020	-0.006	-0.029	3

	0.144; 0.462	0.046; 0.273	0.030; 0.170	-0.016; 0.058	-0.025; 0.014	-0.069; 0.010	
5	0.304	0.163	0.059	0.023	0.001	0.000	3
	0.151; 0.461	0.067; 0.257	-0.019; 0.136	-0.015; 0.060	-0.054; 0.057	-0.021; 0.020	
6	0.398	0.129	0.112	0.022	0.004	-0.055	3
	0.194; 0.602	0.050; 0.209	0.010; 0.214	0.004; 0.039	-0.035; 0.042	-0.111; 0.002	

* CIs were determined by projecting each eigenvector (e_1: e_6 as reported in Table S3) through the 10,000 REML-MVN samples of the respective population's **M**. Although a positive lower CI suggests statistical support for non-zero variance, REML-MVN (and Bayesian MCMC posterior) CI for eigenvalues are known to be biased, leading to false positive inference for leading eigenvalues (Sztepanacz & Blows, 2017). We rely instead on the robust Factor Analytic modelling (Table S2) to infer statistical support.

Table 2. Eigenvalues and eigenvectors of the Krzanowski subspace matrix, H. The 95% CI (below estimate)* for the eigenvalues (λ) of the observed data **H** were determined from Krzanowski subspace analyses applied to each of the 10,000 REML-MVN samples of each of the 12 **M**. For the null data **H**, the reported eigenvalue is the mean (and 95% CI are the 5% quantiles) of the Krzanowski subspace analyses applied to the 12 **M** estimated from the 1000 randomised datasets.

	h_1	h_2	h_3
λ	11.62	8.86	6.62
	10.35; 11.49	6.87; 9.27	5.16; 7.14
Null λ	11.38	10.93	8.56
	10.61; 11.79	9.35; 11.66	5.77; 10.88
CS	-0.37	0.76	-0.32
ILD1.2	0.27	0.41	0.36
ILD1.5	0.65	-0.05	-0.46
ILD2.5	0.20	-0.03	0.64
ILD2.8	0.16	0.42	0.34
ILD3.7	-0.55	-0.27	0.19

* Eigenvalues of **H** are bounded statistics (ranging from 0 to p – here, 0 to 12). In conjunction with the nature of sampling error, which independently affects each estimated **M** and thus contributes dissimilarity, this has resulted in the eigenvalue of h_1 falling above its upper REML-MVN CI. Importantly, the observed λ_1 is within the range observed for the null λ_1, which estimates the error in h_1 when the 12 samples of **M** are all drawn from the same population.

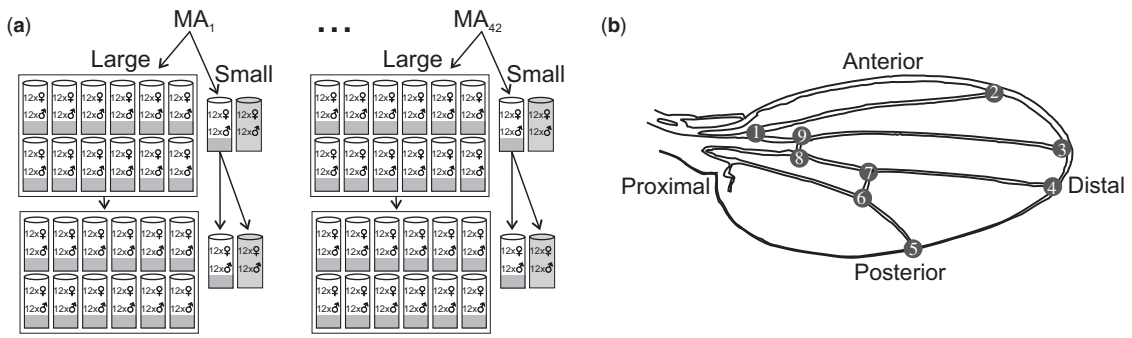

Figure 1. Schematic of (a) experimental design and (b) *Drosophila* wing traits. (a) 42 MA lines each founded two sublines: Small (S; 12 virgin males and 12 virgin females) and Large (L; 144 virgin males and 144 virgin females, distributed evenly among 12 vials). These 84 lines (S and L subline per 42 MA lines) were maintained at these census population sizes for six generations (first two shown here). Each generation, all emergent flies from the 12 vials per L subline were pooled prior to virgin collection. For S sublines, the focal vial contributed offspring to the next generation, while a second, replicate, vial (grey shaded) was used for wing collection only. Each generation, six randomly chosen males were sampled from each of two vials per line (focal and replicate vials for S; randomly chosen two for L). (b) Positions of nine landmarks were recorded and aligned. Six traits were analysed, five inter-landmark distances (named by their end landmarks: ILD1.2, ILD1.5, ILD2.5, ILD2.8, ILD3.7) and wing size, characterised as centroid size (CS).

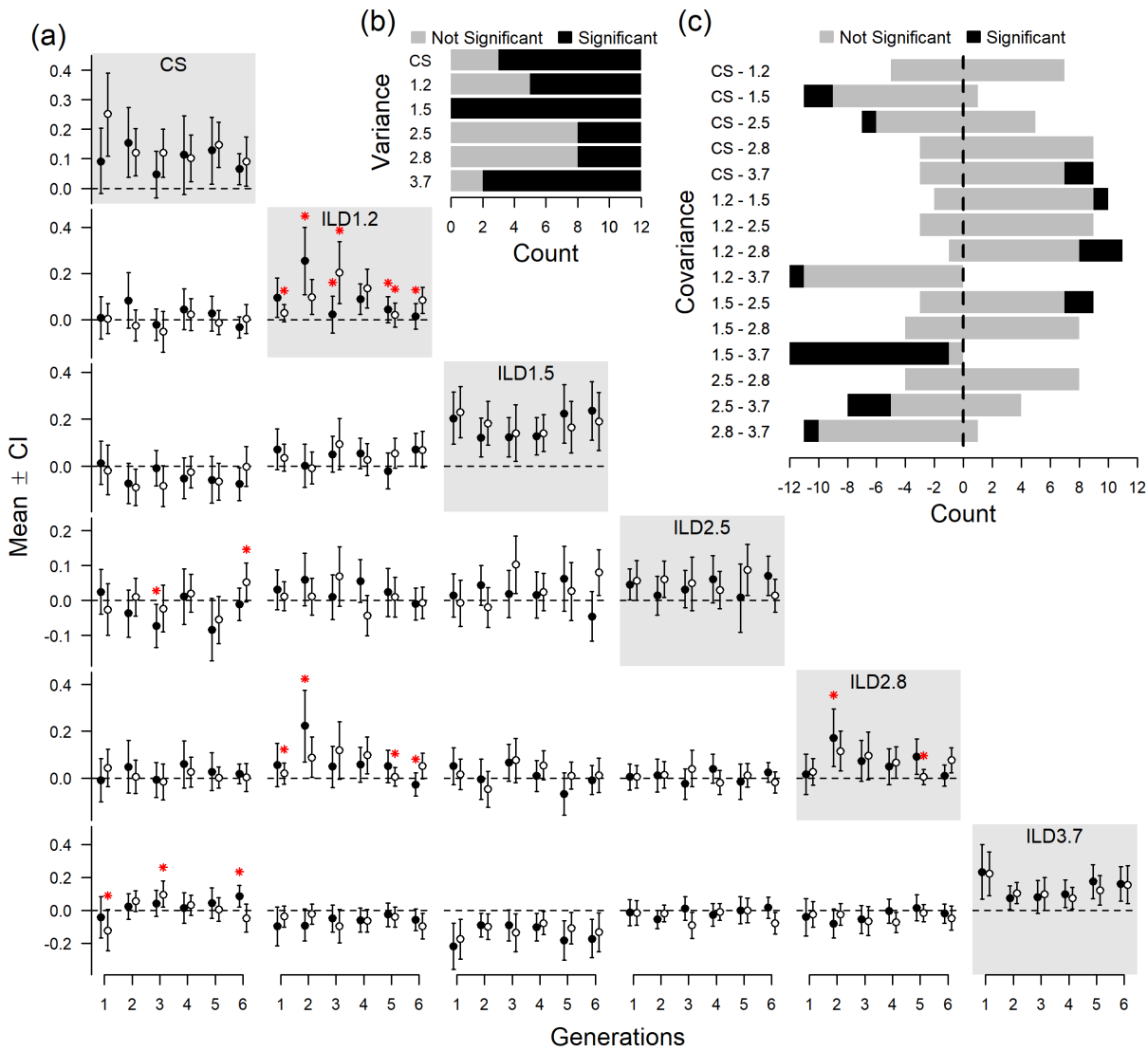

Figure 2. The variability among the 12 population sample estimates of among-line variance and covariance. (a) The REML estimate (point; REML-MVN bootstrap confidence intervals whiskers) of among-line variance (on diagonal, grey boxes; 90% CI) and covariances (below the diagonal; 95% CI) in each population size treatment (Small or Large, closed and open symbols, respectively) in each of six generations. The horizontal dashed line indicates zero; statistical support was concluded when CIs did not include zero (a one-tailed test for variances and two-tailed for covariances). Red stars highlight estimates where CI did not overlap with at least one other estimate of the same parameter. Inset plots are the counts of the 12 estimates of (b) variance for each trait or (c) pairwise covariances between traits that were (black) or were not (grey) statistically supported as different from zero. For the covariances,

the counts of estimates are further plotted by sign (negative or positive). See Figure 1b for trait definitions and Table S1 for the parameter estimates. NB: two estimates of covariance between ILD2.5 and ILD3.7 were exactly 0; these were classified as positive.

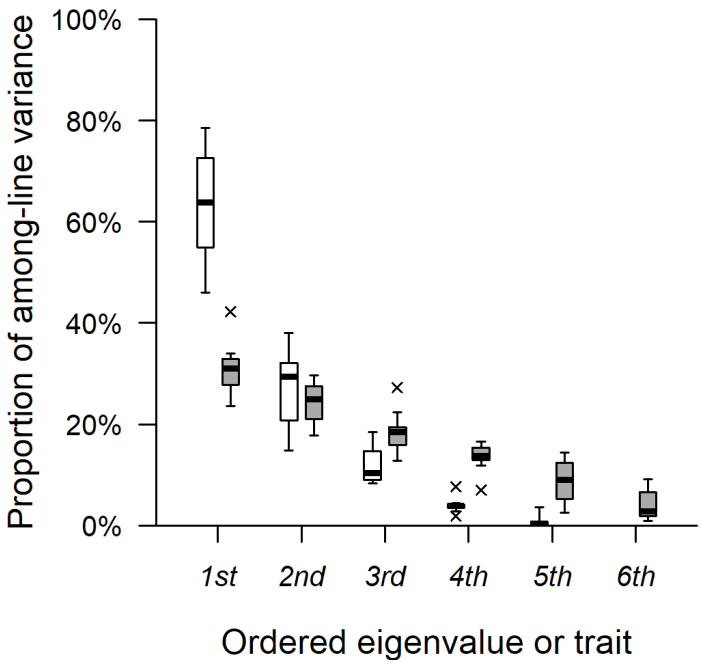

Figure 3. The distribution of among-line variation of six wing traits in 12 population samples. Open boxplots represent distribution across the 12 **M** of the ranked eigenvalues (largest to smallest, x-axis 1st to 6th), plotted as the proportion of the variance in their respective **M**. For comparison, among-line variances of individual traits (diagonals of Table S1 and Figure 2a) were ranked from largest to smallest within each population sample, and their distribution also plotted (grey boxplots). Note, the combination of the six traits (i.e., eigenvector) associated with a specific ordered eigenvalue (1st through 6th), and the specific trait ranked 1st through 6th may differ among the 12 matrices. For each boxplot, the median of the distribution is shown as a black band, boxes represent the interquartile range, whiskers represent 1.5 times the interquartile range, and estimates that fell outside this range are designated with crosses.

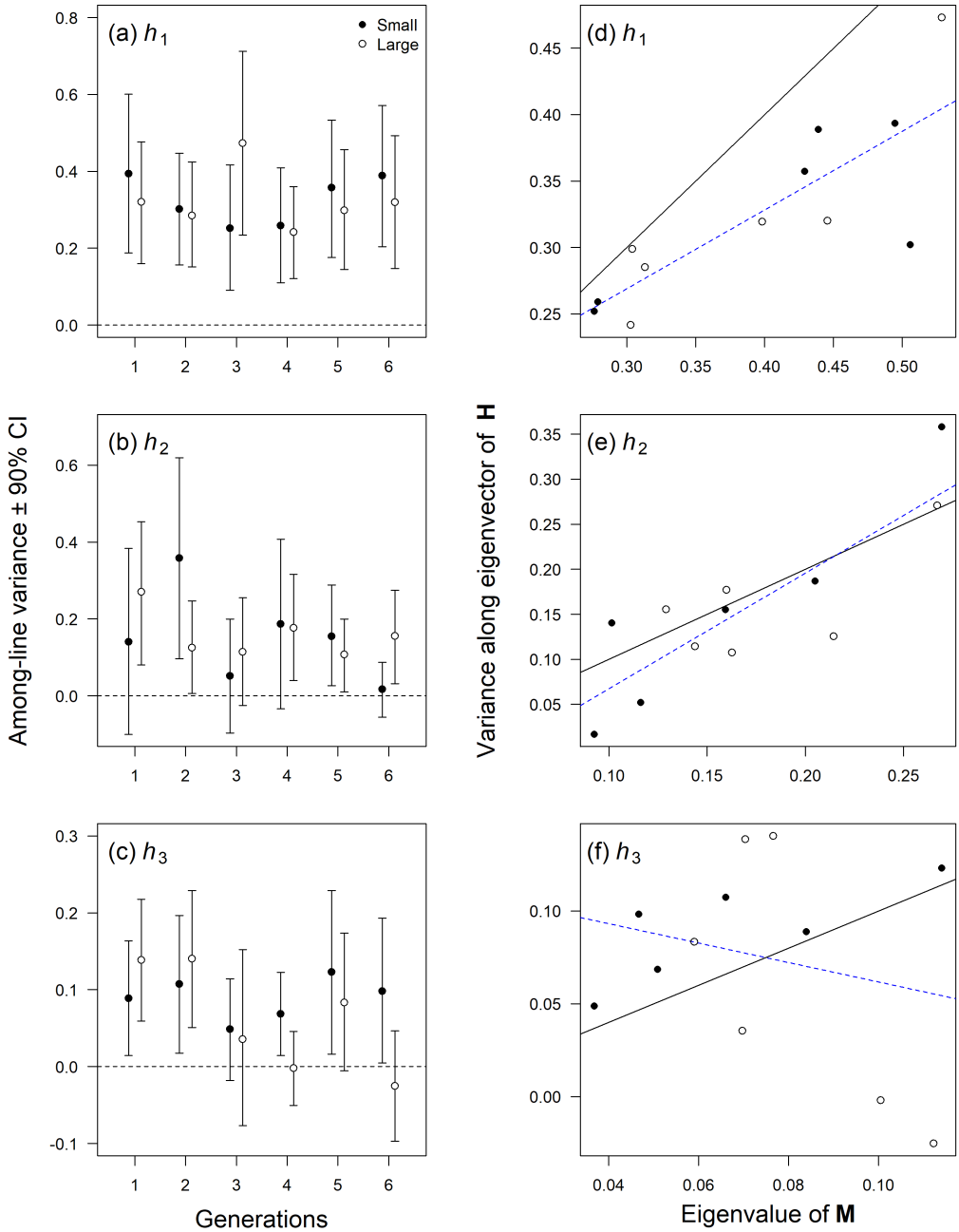

Figure 4. The among-line variance associated with each of the three eigenvectors of the Krzanowski's subspace. Left-hand panels: The magnitude of the among-line variance (90% REML-MVN CI) is plotted for (a) h_1, (b) h_2 and (c) h_3. Within each plot, the populations are ordered by generation (x-axis), then by population size treatments (Small or Large, closed and open symbols, respectively). The dashed lines indicate zero. Note that the y-axis range differs among panels, reflecting decreasing total variance associated with each successive common subspace vector (a to c). Right-hand panels: the magnitude of variance along eigenvectors of

H are plotted against the same-rank eigenvalue of each population sample's respective **M**. (a) h_1 vs m_1, (b) h_2 vs m_2 and (c) h_3 vs m_3. Population size treatment is again indicated by symbol (Small or Large, closed and open symbols, respectively). The dashed blue line indicates the regression slope; solid grey line is the 1:1 line indicating equal magnitude along x and y axes.

SUPPLEMENTARY MATERIALS

Table S1 Mutational variances (diagonal), covariances (lower off-diagonal) and correlations (upper off-diagonal) for the small and large population treatments, for each of the six generations, for six wing traits. Confidence intervals (below variance and covariance estimates) were determined using REML-MVN, and are 90% for variances (for which negative values are not possible: the frequency of negative REML-MNV estimates equate to the LRT probability[1]) and 95% for covariances (reflecting a two-tailed test for a parameter that can be positive or negative). No confidence intervals are reported for correlations because some MVN samples of variances are negative, precluding calculation. To calculate the average absolute strength of correlations, estimates exceeding theoretical bounds (-1 and 1) were set to |1|.

		Small Population						Large Population					
Gen		**CS**	**1.2**	**1.5**	**2.5**	**2.8**	**3.7**	**CS**	**1.2**	**1.5**	**2.5**	**2.8**	**3.7**
1	**CS**	0.091	0.096	0.096	0.375	-0.236	-0.288	0.252	0.058	-0.075	-0.217	0.533	-0.519
		-0.018; 0.204						0.109; 0.390					
	1.2	0.009	0.096	0.516	0.472	1.403	-0.649	0.005	0.029	0.441	0.295	0.679	-0.473
		-0.083; 0.100	0.011; 0.181					-0.059; 0.070	-0.008; 0.066				
	1.5	0.013	0.072	0.203	0.157	0.895	-1.012	-0.018	0.036	0.230	-0.061	0.190	-0.768
		-0.078; 0.106	-0.014; 0.159	0.094; 0.315				-0.122; 0.091	-0.021; 0.094	0.121; 0.339			
	2.5	0.024	0.031	0.015	0.045	0.186	-0.127	-0.026	0.012	-0.007	0.057	0.153	-0.133
		-0.039; 0.089	-0.027; 0.088	-0.047; 0.077	0.002; 0.090			-0.099; 0.048	-0.029; 0.054	-0.074; 0.058	-0.000; 0.114		
	2.8	-0.009	0.055	0.051	0.005	0.016	-0.671	0.044	0.019	0.015	0.006	0.027	-0.335
		-0.102; 0.084	-0.037; 0.148	-0.028; 0.129	-0.051; 0.061	-0.070; 0.102		-0.036; 0.123	-0.027; 0.064	-0.050; 0.080	-0.043; 0.054	-0.031; 0.083	
	3.7	-0.042	-0.097	-0.22	-0.013	-0.041	0.233	-0.123	-0.038	-0.174	-0.015	-0.026	0.223
		-0.169; 0.082	-0.217; 0.021	-0.359; -0.079	-0.092; 0.065	-0.155; 0.072	0.068; 0.400	-0.246; 0.006	-0.103; 0.026	-0.295; -0.056	-0.089; 0.059	-0.103; 0.053	0.090; 0.353
2	**CS**	0.154	0.420	-0.533	-0.749	0.289	0.205	0.121	-0.230	-0.606	0.116	0.043	0.497
		0.038; 0.273						0.042; 0.203					
	1.2	0.083	0.254	0.011	0.972	1.072	-0.667	-0.025	0.098	-0.060	0.155	0.833	-0.227
		-0.036; 0.204	0.108; 0.399					-0.091; 0.042	0.023; 0.175				
	1.5	-0.073	0.002	0.122	1.029	-0.035	-0.951	-0.09	-0.008	0.182	-0.190	-0.333	-0.716
		-0.158; 0.013	-0.089; 0.095	0.040; 0.205				-0.167; -0.013	-0.076; 0.061	0.091; 0.277			

	2.5	-0.036	0.06	0.044	0.015	0.236	-1.670	0.01	0.012	-0.02	0.061	0.156	-0.237
		-0.105; 0.030	-0.015; 0.135	-0.013; 0.100	-0.042; 0.070			-0.045; 0.064	-0.042; 0.064	-0.077; 0.037	0.009; 0.113		
	2.8	0.047	0.224	-0.005	0.012	0.172	-0.713	0.005	0.088	-0.048	0.013	0.114	-0.229
		-0.064; 0.160	0.068; 0.376	-0.091; 0.081	-0.060; 0.082	0.049; 0.296		-0.066; 0.077	0.004; 0.175	-0.125; 0.029	-0.043; 0.070	0.030; 0.200	
	3.7	0.022	-0.092	-0.091	-0.056	-0.081	0.075	0.056	-0.023	-0.099	-0.019	-0.025	0.105
		-0.054; 0.100	-0.187; 0.007	-0.163; -0.018	-0.111; -0.001	-0.167; 0.008	0.006; 0.147	-0.007; 0.120	-0.085; 0.039	-0.178; -0.022	-0.070; 0.031	-0.090; 0.040	0.041; 0.170
3	**CS**	0.047	-0.639	-0.118	-1.882	-0.137	0.685	0.12	-0.333	-0.646	-0.313	-0.149	0.890
		-0.033; 0.125						0.037; 0.201					
	1.2	-0.021	0.023	0.959	0.369	1.196	-1.142	-0.052	0.203	0.556	0.692	0.845	-0.684
		-0.090; 0.046	-0.058; 0.102					-0.141; 0.037	0.069; 0.337				
	1.5	-0.009	0.051	0.123	0.303	0.697	-0.917	-0.084	0.094	0.141	1.239	0.653	-1.163
		-0.084; 0.066	-0.025; 0.128	0.040; 0.208				-0.172; 0.002	-0.015; 0.204	0.021; 0.261			
	2.5	-0.073	0.01	0.019	0.032	-0.517	0.217	-0.024	0.069	0.103	0.049	0.583	-1.320
		-0.135; -0.011	-0.054; 0.073	-0.049; 0.087	-0.021; 0.086			-0.090; 0.044	-0.016; 0.154	0.020; 0.185	-0.029; 0.125		
	2.8	-0.008	0.049	0.066	-0.025	0.073	-0.733	-0.016	0.118	0.076	0.04	0.096	-0.674
		-0.083; 0.067	-0.040; 0.136	-0.014; 0.144	-0.091; 0.040	-0.014; 0.161		-0.094; 0.060	-0.006; 0.240	-0.020; 0.170	-0.037; 0.119	-0.007; 0.197	
	3.7	0.042	-0.049	-0.091	0.011	-0.056	0.080	0.096	-0.096	-0.136	-0.091	-0.065	0.097
		-0.037; 0.121	-0.131; 0.032	-0.185; 0.003	-0.061; 0.084	-0.141; 0.029	-0.019; 0.181	0.019; 0.177	-0.199; 0.006	-0.253; -0.023	-0.170; -0.014	-0.154; 0.023	0.000; 0.199
4	**CS**	0.114	0.447	-0.432	0.132	0.789	0.132	0.102	0.196	-0.200	0.362	0.317	0.357
		-0.020; 0.245						0.023; 0.180					
	1.2	0.045	0.089	0.517	0.760	0.878	-0.639	0.023	0.135	0.203	-0.676	1.028	-0.640
		-0.043; 0.134	0.023; 0.155					-0.047; 0.092	0.050; 0.218				
	1.5	-0.052	0.055	0.127	0.182	0.114	-0.919	-0.024	0.028	0.141	0.384	0.560	-0.754
		-0.139; 0.035	-0.011; 0.119	0.049; 0.208				-0.091; 0.041	-0.039; 0.097	0.062; 0.220			
	2.5	0.011	0.056	0.016	0.061	0.732	-0.360	0.02	-0.043	0.025	0.030	-0.427	-0.212
		-0.068; 0.091	-0.005; 0.117	-0.050; 0.082	-0.007; 0.129			-0.034; 0.075	-0.101; 0.014	-0.029; 0.078	-0.024; 0.084		
	2.8	0.059	0.058	0.009	0.04	0.049	-0.072	0.026	0.097	0.054	-0.019	0.066	-1.045
		-0.044; 0.159	-0.018; 0.132	-0.057; 0.074	-0.023; 0.103	-0.028; 0.125		-0.038; 0.090	0.017; 0.176	-0.008; 0.117	-0.071; 0.033	-0.002; 0.133	
	3.7	0.014	-0.06	-0.103	-0.028	-0.005	0.099	0.031	-0.064	-0.077	-0.010	-0.073	0.074
		-0.079; 0.106	-0.133; 0.015	-0.187; -0.018	-0.096; 0.040	-0.077; 0.068	0.016; 0.184	-0.031; 0.093	-0.132; 0.004	-0.148; -0.006	-0.060; 0.041	-0.134; -0.012	0.008; 0.140

5	**CS**	0.129	0.358	-0.341	-2.465	0.249	0.287	0.147	-0.216	-0.417	-0.484	0.074	0.030
		0.015; 0.241						0.070; 0.224					
	1.2	0.027	0.044	-0.201	1.206	0.806	-0.297	-0.012	0.021	0.934	0.233	0.488	-0.787
		-0.048; 0.101	-0.012; 0.100					-0.064; 0.040	-0.031; 0.073				
	1.5	-0.058	-0.020	0.224	1.403	-0.476	-0.927	-0.065	0.055	0.165	0.224	0.348	-0.765
		-0.157; 0.041	-0.097; 0.057	0.098; 0.348				-0.144; 0.013	-0.008; 0.120	0.056; 0.277			
	2.5	-0.084	0.024	0.063	0.009	-0.524	0.000	-0.055	0.01	0.027	0.088	0.572	0.000
		-0.173; 0.006	-0.046; 0.092	-0.030; 0.155	-0.091; 0.105			-0.124; 0.011	-0.048; 0.067	-0.057; 0.109	0.015; 0.161		
	2.8	0.027	0.051	-0.068	-0.015	0.091	0.111	0.002	0.005	0.01	0.012	0.005	-0.686
		-0.054; 0.109	-0.019; 0.119	-0.159; 0.022	-0.091; 0.060	0.016; 0.167		-0.044; 0.048	-0.035; 0.046	-0.047; 0.069	-0.039; 0.063	-0.028; 0.037	
	3.7	0.043	-0.026	-0.183	0	0.014	0.174	0.004	-0.04	-0.109	0	-0.017	0.123
		-0.048; 0.136	-0.099; 0.045	-0.303; -0.062	-0.081; 0.082	-0.063; 0.094	0.071; 0.278	-0.065; 0.077	-0.101; 0.021	-0.205; -0.012	-0.075; 0.073	-0.068; 0.036	0.032; 0.211
6	**CS**	0.066	-1.086	-0.601	-0.161	0.662	0.825	0.091	0.046	-0.015	1.485	0.036	-0.404
		0.013; 0.118						0.008; 0.174					
	1.2	-0.033	0.014	1.235	-0.317	-2.366	-1.201	0.004	0.084	0.554	-0.204	0.634	-0.833
		-0.078; 0.013	-0.041; 0.069					-0.059; 0.065	0.028; 0.141				
	1.5	-0.075	0.071	0.236	-0.355	-0.185	-0.888	-0.002	0.07	0.190	1.551	0.091	-0.775
		-0.146; -0.006	0.001; 0.141	0.110; 0.360				-0.086; 0.083	-0.007; 0.148	0.067; 0.313			
	2.5	-0.011	-0.01	-0.046	0.071	0.901	0.150	0.053	-0.007	0.08	0.014	-0.548	-1.696
		-0.057; 0.035	-0.056; 0.036	-0.117; 0.026	0.015; 0.127			-0.004; 0.107	-0.052; 0.038	0.014; 0.146	-0.034; 0.061		
	2.8	0.017	-0.028	-0.009	0.024	0.010	-0.498	0.003	0.051	0.011	-0.018	0.077	-0.439
		-0.025; 0.059	-0.077; 0.022	-0.070; 0.053	-0.018; 0.066	-0.035; 0.056		-0.058; 0.063	-0.004; 0.106	-0.061; 0.086	-0.064; 0.026	0.023; 0.130	
	3.7	0.085	-0.057	-0.173	0.016	-0.020	0.161	-0.048	-0.095	-0.133	-0.079	-0.048	0.155
		0.019; 0.151	-0.122; 0.009	-0.288; -0.055	-0.050; 0.081	-0.079; 0.038	0.055; 0.264	-0.132; 0.037	-0.174; -0.018	-0.253; -0.015	-0.145; -0.013	-0.120; 0.025	0.040; 0.271

[1.] Dugand RJ, Aguirre JD, Hine E, Blows MW, McGuigan K. 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 USA. 2021;118(31): e2026217118.

Table S2. Factor analytic modelling of among-line covariance in each of the 12 populations. Sequential REML mixed models (model 1) were fit with zero up to six dimensions of among-line variance (columns, left to right; the number of among-line parameters estimated by the model is shown in parentheses on the second row). The -2 loglikelihood (-2LL) of each model is reported. Statistical support for each additional dimension was determined by log-likelihood ratio tests (*P*-values shown below the -2LL for that dimension); the difference in -2LL between sequential models follows a chi-squared distribution with degrees of freedom determined by the difference in number of parameters. *P*-values are shown in bold where the null hypothesis (variance along that dimension is not greater than zero) was rejected.

Treat	**Gen**	None (0)	1 (6)	2 (11)	3 (15)	4 (18)	5 (20)	6 (21)
	1	6218.8	6194.3 ***0.0004***	6184.3 *0.0752*	6176.4 *0.0953*	6175.6 *0.8495*	6175.6 *>0.9999*	6175.6 *>0.9999*
	2	6474.0	6444.6 ***0.0001***	6430.7 ***0.0163***	6428.0 *0.6092*	6427.1 *0.8254*	6427.1 *>0.9999*	6427.1 *>0.9999*
Small	3	5676.5	5660.3 ***0.0127***	5647.4 ***0.0243***	5646.7 *0.9513*	5646.7 *>0.9999*	5646.1 *0.7408*	5646.2 *0.7518*
	4	6603.6	6588.6 ***0.0203***	6577.4 ***0.0476***	6575.6 *0.7725*	6574.5 *0.7771*	6574.5 *>0.9999*	6574.5 *>0.9999*
	5	5489.5	5466.3 ***0.0007***	5460.4 *0.3161*	5456.1 *0.3669*	5455.9 *0.9776*	5452.7 *0.2019*	5452.7 *>0.9999*

	6	6633.6	6590.0 ***<0.0001***	6578.5 ***0.0423***	6573.9 *0.3309*	6572.1 *0.6149*	6572.1 *>0.9999*	6572.1 *>0.9999*
Large	1	6473.9	6448.4 ***0.0003***	6429.6 ***0.0021***	6420.4 *0.0563*	6417.2 *0.3618*	6416.5 *0.7047*	6416.5 *>0.9999*
	2	6833.0	6795.8 ***<0.0001***	6785.8 *0.0752*	6775.9 ***0.0421****	6771.0 *0.1793*	6769.8 *0.5488*	6769.8 *>0.9999*
	3	5498.8	5477.7 ***0.0018***	5457.5 ***0.0011***	5453.7 *0.4337*	5452.9 *0.8495*	5452.9 *>0.9999*	5452.8 *0.7518*
	4	6415.9	6390.3 ***0.0003***	6377.7 ***0.0274***	6371.6 *0.1918*	6371.5 *0.9918*	6371.5 *>0.9999*	6370.7 *0.3711*
	5	6047.7	6032.6 ***0.0195***	5998.9 ***<0.0001***	5994.9 *0.4060*	5994.0 *0.8254*	5994.0 *>0.9999*	5994.0 *>0.9999*
	6	5455.0	5435.1 ***0.0029***	5424.6 *0.0622*	5417.4 *0.1257*	5413.2 *0.2407*	5413.2 *>0.9999*	5413.1 *0.7518*

* statistical support for three but not two dimensions suggests re-ordering such that the FA0(2) model estimated a lower variance subset of principal components (MEYER AND KIRKPATRICK 2008). For this population sample, we interpret two dimensions as being statistically supported.

Meyer, K., and M. Kirkpatrick. Perils of parsimony: Properties of reduced rank estimates of genetic covariance matrices. Genetics 2008; 180**:** 1153-1166.

Table S3. The eigenvectors for M. Eigenanalyses were performed separately for **M** in each population (small, S, or large, L for every generation). The corresponding eigenvalues are reported in Table 2.

Population	Trait	m_1	m_2	m_3	m_4	m_5	m_6
S1	**size**	0.10	0.90	0.14	0.26	-0.28	0.06
	1.2	0.31	0.01	-0.77	0.33	0.16	0.42
	1.5	0.63	-0.20	0.26	-0.31	-0.47	0.43
	2.5	0.07	0.34	-0.40	-0.82	0.11	-0.18
	2.8	0.16	-0.16	-0.34	0.17	-0.61	-0.65
	3.7	-0.68	-0.05	-0.21	-0.14	-0.53	0.42
S2	**size**	0.21	0.60	0.75	0.00	-0.07	-0.18
	1.2	0.72	0.04	-0.09	-0.37	0.37	0.45
	1.5	0.03	-0.63	0.52	0.27	0.50	-0.10
	2.5	0.12	-0.27	0.02	-0.7	-0.13	-0.64
	2.8	0.59	0.03	-0.32	0.55	-0.08	-0.50
	3.7	-0.28	0.41	-0.26	-0.06	0.76	-0.32
S3	**size**	-0.18	0.68	-0.34	0.08	0.21	-0.58
	1.2	0.33	-0.03	0.22	0.22	0.89	0.08
	1.5	0.63	0.06	-0.7	0.16	-0.12	0.27
	2.5	0.05	-0.67	-0.18	0.32	-0.02	-0.65
	2.8	0.43	0.31	0.53	0.51	-0.39	-0.16
	3.7	-0.53	0.00	-0.16	0.74	0.02	0.37
S4	**size**	-0.01	0.70	0.54	-0.13	0.21	0.40
	1.2	0.49	0.34	-0.09	0.07	-0.8	-0.02
	1.5	0.59	-0.36	0.19	0.49	0.21	0.45
	2.5	0.28	0.23	-0.75	-0.32	0.32	0.33
	2.8	0.20	0.42	-0.12	0.51	0.37	-0.60
	3.7	-0.55	0.17	-0.31	0.61	-0.18	0.41
S5	**size**	-0.29	0.64	0.52	-0.32	0.12	0.34
	1.2	-0.04	0.35	-0.44	-0.55	-0.49	-0.37
	1.5	0.72	0.10	0.18	-0.33	0.48	-0.32
	2.5	0.17	-0.30	-0.32	-0.49	0.07	0.73
	2.8	-0.21	0.39	-0.62	0.21	0.61	0.02
	3.7	-0.57	-0.46	0.11	-0.44	0.37	-0.34
S6	**size**	-0.29	0.36	0.62	0.29	-0.52	0.23
	1.2	0.22	0.05	-0.35	0.16	-0.68	-0.58
	1.5	0.71	0.23	0.19	0.57	0.29	0.00
	2.5	-0.11	-0.81	0.00	0.53	-0.12	0.17
	2.8	-0.02	-0.28	0.64	-0.17	0.18	-0.67
	3.7	-0.59	0.28	-0.24	0.51	0.37	-0.35

Table S3 (cont.)

Population	Trait	m_1	m_2	m_3	m_4	m_5	m_6
L1	**size**	0.42	0.81	0.09	0.02	-0.09	0.39
	1.2	0.12	-0.07	-0.28	0.60	0.71	0.21
	1.5	0.55	-0.57	0.28	0.13	-0.29	0.43
	2.5	-0.01	-0.08	-0.87	-0.19	-0.28	0.33
	2.8	0.11	0.10	-0.17	0.69	-0.52	-0.46
	3.7	-0.70	0.04	0.21	0.34	-0.24	0.54
L2	**size**	-0.49	0.13	-0.52	0.64	0.25	-0.07
	1.2	0.03	-0.63	0.22	0.04	0.62	-0.41
	1.5	0.73	0.14	-0.08	0.35	0.39	0.41
	2.5	-0.05	-0.15	-0.66	-0.63	0.29	0.26
	2.8	-0.12	-0.70	0.00	0.24	-0.33	0.57
	3.7	-0.46	0.25	0.49	-0.12	0.45	0.52
L3	**size**	-0.3	0.54	-0.69	0.15	0.01	-0.34
	1.2	0.51	0.60	0.39	0.45	0.15	-0.06
	1.5	0.49	-0.28	-0.40	-0.09	0.71	-0.12
	2.5	0.31	-0.08	-0.46	0.36	-0.35	0.67
	2.8	0.33	0.43	-0.08	-0.80	-0.17	0.18
	3.7	-0.45	0.27	0.10	-0.04	0.57	0.62
L4	**size**	-0.01	0.55	0.76	-0.05	-0.11	0.34
	1.2	0.56	0.48	-0.24	-0.13	0.62	-0.04
	1.5	0.48	-0.58	0.37	-0.55	0.05	0.06
	2.5	-0.06	-0.23	0.48	0.51	0.5	-0.45
	2.8	0.49	0.20	0.06	0.13	-0.58	-0.6
	3.7	-0.47	0.21	0.07	-0.64	0.12	-0.55
L5	**size**	-0.39	0.72	-0.45	-0.32	0.11	-0.04
	1.2	0.23	0.11	-0.07	-0.23	-0.92	-0.15
	1.5	0.71	0.16	0.11	-0.59	0.31	0.13
	2.5	0.20	-0.45	-0.84	-0.03	0.10	-0.18
	2.8	0.06	0.04	-0.23	0.17	-0.16	0.94
	3.7	-0.49	-0.48	0.11	-0.69	-0.05	0.19
L6	**size**	0.14	0.48	0.70	0.00	0.46	0.22
	1.2	0.35	-0.49	0.13	-0.63	0.35	-0.32
	1.5	0.62	0.13	-0.56	0.26	0.42	0.20
	2.5	0.26	0.43	0.02	0.14	-0.17	-0.84
	2.8	0.16	-0.58	0.36	0.7	0.06	-0.13
	3.7	-0.62	0.00	-0.22	0.14	0.67	-0.31

Table S4. Pairwise dot products between eigenvectors of M **and the eigenvectors of** H. The largest dot product (indicating the strongest relationship) is bolded for each population sample's comparison for each **H** vector. Dot products of normalised vectors range in magnitude from 0 (orthogonal orientation) to 1 (identical orientation). Dot products can be positive or negative, but we focus here on the absolute magnitude for two reasons. First, eigenvector loadings have arbitrary direction from the origin along the defined axis such that the sign of all loadings can be flipped (positive to negative and negative to positive) with no effect on the interpretation of trait covariation. Second, eigenvectors of **H** are not constrained to be orthogonal within the space of each **M**, which can result in opposing sign of association with two orthogonal eigenvectors of **M** if the **H** vector falls between them. The eigenvectors in Table S2 were flipped such that the focal dot products (i.e., between same rank **H** and **M** eigenvectors) were typically positive. Note that no m_5 or m_6 eigenvectors from any population sample were included in the common subspace analysis, and m_4 was included for only one matrix (Table 1). Dot products (r) can be converted to angle degree by: $(cos^{-1}r)(^{180}/_{\pi})$.

Population	**H** vector	m_1	m_2	m_3	m_4	m_5	m_6
S1	h_1	**0.87**	-0.39	-0.11	-0.27	0.06	0.00
S2	h_1	0.41	**-0.90**	0.13	0.06	-0.01	0.09
S3	h_1	**0.94**	-0.31	-0.13	-0.13	0.01	0.05
S4	h_1	**0.91**	-0.38	-0.10	0.07	0.07	-0.12
S5	h_1	**0.88**	0.18	-0.42	-0.07	0.04	-0.10
S6	h_1	**0.93**	-0.33	0.03	0.10	0.00	-0.12
L1	h_1	0.64	**-0.71**	-0.24	0.12	0.03	-0.11
L2	h_1	**0.89**	-0.41	-0.20	-0.02	0.09	0.04
L3	h_1	**0.93**	-0.32	-0.06	-0.03	0.09	-0.15
L4	h_1	**0.79**	-0.58	-0.04	0.10	0.18	0.02
L5	h_1	**0.99**	0.05	-0.05	0.07	-0.07	0.07
L6	h_1	**0.86**	-0.23	-0.41	0.06	-0.20	-0.06
S1	h_2	0.42	**0.63**	-0.30	0.48	-0.24	-0.18
S2	h_2	**0.77**	0.41	0.44	0.10	-0.16	-0.05
S3	h_2	0.29	**0.65**	0.14	0.15	0.36	-0.57
S4	h_2	0.39	**0.81**	0.42	-0.04	0.02	-0.10
S5	h_2	-0.21	**0.92**	-0.07	-0.23	0.02	0.20

S6	h_2	-0.01	0.11	0.65	0.03	**-0.71**	-0.26
L1	h_2	0.58	**0.65**	-0.16	0.46	0.09	0.01
L2	h_2	-0.32	-0.52	-0.41	**0.64**	0.16	-0.15
L3	h_2	0.21	**0.78**	-0.39	-0.03	-0.18	-0.39
L4	h_2	0.53	**0.68**	0.45	0.15	-0.12	0.15
L5	h_2	-0.09	**0.74**	-0.48	-0.05	-0.37	0.25
L6	h_2	0.45	-0.10	**0.82**	-0.02	0.32	0.08
S1	h_3	0.24	0.04	**0.85**	0.32	-0.13	0.32
S2	h_3	-0.40	-0.03	**0.66**	0.53	0.04	0.37
S3	h_3	0.04	**0.58**	-0.54	-0.50	-0.17	0.31
S4	h_3	-0.05	-0.39	**0.87**	0.07	0.16	0.26
S5	h_3	0.32	0.27	**0.80**	0.27	0.11	-0.31
S6	h_3	0.34	**0.76**	0.24	-0.08	0.16	0.47
L1	h_3	0.45	0.03	**0.83**	-0.33	-0.02	0.09
L2	h_3	0.33	0.62	0.05	**0.70**	-0.12	-0.15
L3	h_3	-0.28	-0.32	-0.24	-0.11	0.45	-0.75
L4	h_3	-0.02	-0.22	0.16	-0.47	-0.38	**0.75**
L5	h_3	0.06	**0.63**	0.53	-0.20	0.51	-0.14
L6	h_3	0.10	0.31	-0.17	-0.01	0.18	**0.92**