Dis-integrating the fly: A mutational perspective on phenotypic integration and covariation.

Running title: Mutational effects on environmental covariation

Authors: Annat Haber[1,2] & Ian Dworkin[1,3,4]

Author affiliations:

1) BEACON Center for the study of Evolution in Action. Michigan State University.

2) Department of Zoology, Tel Aviv University

3) Department of Integrative Biology, Michigan State University.

4) Department of Biology, McMaster University.

Correspondence: annat22@gmail.com, dworkin@mcmaster.ca

Key words: Drosophila, integration, canalization, morphometrics, covariation, mutation

Number of words: 6755

Number of tables: 3

Number of figures: 7

Data will be archived in DRYAD

Abstract

The structure of environmentally induced phenotypic covariation can influence the effective strength and magnitude of natural selection. Yet our understanding of the factors that contribute to and influence the evolutionary lability of such covariation is poor. Most studies have either examined environmental variation without accounting for covariation, or examined phenotypic and genetic covariation without distinguishing the environmental component. In this study we examined the effect of mutational perturbations on different properties of environmental covariation, as well as mean shape. We use strains of *Drosophila melanogaster* bearing well-characterized mutations known to influence wing shape, as well as naturally-derived strains, all reared under carefully-controlled conditions and with the same genetic background. We find that mean shape changes more freely than the covariance structure, and that different properties of the covariance matrix change independently from each other. The perturbations affect matrix orientation more than they affect matrix eccentricity or total variance. Yet, mutational effects on matrix orientation do not cluster according to the developmental pathway that they target. These results suggest that it might be useful to consider a more general concept of 'decanalization', involving all aspects of variation and covariation.

Introduction

The evolutionary response to natural selection requires phenotypic variation. As such, the mechanisms generating phenotypic variation and covariation (**P**) are of fundamental importance. Most studies focus on the genetic component in **P**, summarized as the genetic covariance matrix **G**, due to its role in the response to natural selection as summarized in the multivariate version of the breeder's equation (Lande 1979; Lande and Arnold 1983). Thus, the generation of heritable components of variation, governed by mutation and recombination, is broadly studied. However, non-heritable sources of phenotypic (co)variation, often summarized through **E**, remain important for our understanding of both the magnitude and direction of response to natural selection through its contributions to **P** (i.e., **P**=**G**+**E**). The basic measure of the strength of selection - the selection differential - depends on **P** via the covariance between the phenotype and fitness (the Robertson-Price identity; Robertson 1966; Price 1970). In addition to its influence on the magnitude of selection, **E** influences the response to selection via **P** as summarized in Lande's (1979) equation ($\Delta Z=\mathbf{G}\mathbf{P}^{-1}s$). Empirical evidence shows that E is shaped by genetically-defined reaction norms to differences in environment, and can thus be adaptive and amenable to directional and stabilizing selection (Hill and Mulder 2010). In a multivariate phenotype, the covariances in **E** affect the orientation of the response to selection in addition to its magnitude. Even in cases when **E** does not facilitate adaptive evolution directly, it can certainly impede it and possibly bias its trajectory towards a secondary peak (Burger 1986). Yet, our understanding of how this environmental covariation is generated during development is incomplete, as is our understanding of how it evolves.

53 Mutational and environmental perturbations have often been utilized to examine how
54 the environmental component of phenotypic variation changes, particularly in the context of
55 canalization (Waddington 1952; Dworkin 2005a,b; Levy and Siegal 2008; Hallgrímsson et al. 2009;
56 Hill and Mulder 2010; Paaby and Rockman 2014; and references therein). A number of these
57 studies have found an increase in phenotypic variance in response to perturbations, often
58 attributed to decanalization, and due in part to a release of cryptic genetic variation (Dworkin
59 2005a,b; Paaby and Rockman 2014; and references therein). Yet, both increases and decreases of
60 environmental variance have also been observed without the release of genetic variation (Debat
61 et al. 2006; Dworkin and Gibson 2006; Debat et al. 2009; Hallgrímsson et al. 2009). Most studies,
62 however, have examined the effect of mutational perturbations on the variance of a trait or on
63 the total variance across several traits (e.g., Dworkin 2005a; Levy and Siegal 2008), whereas in
64 multivariate systems other aspects of covariation need to be considered.
65 Covariance matrices are commonly characterized based on their size, shape, and
66 orientation (Jones et al. 2003; Hohenlohe and Arnold 2008), as illustrated in Fig. 1. Matrix size is
67 often measured as the total variance in this context (sum of the individual trait variances) and it
68 quantifies variation overall, irrespective of how it is distributed among different directions (Fig.
69 1A; henceforth referred to as total variance). The orientation of the matrix refers to the direction
70 in which the primary axes of covariation are pointing (Fig. 1B), and is determined by the relative
71 contribution of each trait to each eigenvector. Thus, a change in matrix orientation reflects a
72 change in the direction in which most of the variation is concentrated, and therefore a change in
73 the combination of traits that are more easily attainable through evolution. The eccentricity of
74 the matrix describes its shape – how much it deviates from a hypersphere – and it quantifies how

evenly the variation is distributed among the different directions (Fig. 1C), irrespective of orientation and total variance. A more eccentric matrix (i.e. more cigar-shaped) means that there is considerably more variation along the first axis relative to other axes, and therefore a fewer trait combinations that are easily attainable. To the extent that the observed covariation reflects integrating factors, higher eccentricity reflects a more integrated body plan.

Previous studies provide mixed expectations as to how the different properties of the covariance matrix might change relative to each other and relative to mean shape. Moreover, the association between changes to matrix orientation and eccentricity has mostly been studied so far for **P** and **G** rather than **E**. Simulations by Jones et al. (2003, 2004, 2007, 2012) and Revell (2007) suggest that the orientation of **G** is more likely to change than eccentricity under most conditions, including effects due to population size, magnitude, and orientation of mutational correlation and stabilizing selection, and different modes of directional selection. In addition, higher eccentricity enhances the stability of the orientation under most conditions. The magnitude and stability of eccentricity and total variance, on the other hand, are mostly influenced by population size rather than mutational correlations and selection. Empirical studies have found evidence that both eccentricity and orientation can be either stable or labile across populations and species at different phylogenetic scales (Arnold et al. 2008; de Oliveira et al. 2009; Porto et al. 2009; Haber 2015a,b). Unfortunately it is not clear to what extent the expectations from these studies (examining **G** and **P**) are relevant to **E**. Changes in **G** reflect a combination of influences from mutation and other evolutionary forces. The structure of **E**, on the other hand, reflects the interaction between the environment and development, and different genotypes differ in their sensitivity to the environment. Thus, while strong directional

97 selection can deplete genetic covariation (Homrigh et al. 2007, Walsh and Blows 2009), altering
98 matrix properties of **G**, it is not yet clear how this should alter the matrix properties of **E** as well.
99 Theory suggests that strong selection, both disruptive and directional, can lead to an
100 increase in the total variance of **E** (Hill and Mulder 2010). Hallgrímsson et al. (2009) postulated
101 that when the total variance increases due to developmental perturbations, matrix eccentricity
102 could either increase or decrease, depending on how the perturbed developmental process
103 contributes to the original patterns of covariation. Eccentricity would increase with total variance
104 if the targeted pathway plays a major role in integrating the measured traits, thus adding
105 variance to the major axis of covariation. When the perturbed pathway does not contribute
106 much to the major axis of covariation, changes in variance are distributed among the smaller
107 subsequent eigenvalues, thus reducing eccentricity. Indeed, empirical studies have found both
108 increases and decreases of eccentricity with an increase in the total variance of **E** (Hallgrímsson
109 et al. 2009). In addition, studies have found mutational effects on matrix orientation of **E** that
110 could not be readily associated with developmental pathways or mutational effect size
111 (Hallgrímsson et al. 2006; Debat et al. 2009; Jamniczky and Hallgrímsson 2009). Most studies,
112 however, did not include a direct comparison of the different properties of **E** across a wide range
113 of mutations, nor how these properties change relative to mean shape. Thus, we are left with
114 little understanding of the lability of **E**.
115 *Drosophila* wing shape is an ideal system to address these questions, and in particular
116 how mutational perturbations influence the different aspects of the covariance structure of **E**.
117 The development of the wing of *D. melanogaster* has been a model system for over 70 years, and
118 it is one of the best genetically characterized model systems. The extensive set of genetic tools,

119 mutant lines, and ability to generate many genetically identical individuals reared under common
120 environments enables the study of mutational effects on **E**. Variation in both size and shape of
121 the *Drosophila* wing has a long history of research, and sophisticated high dimensional
122 representations of shape can be generated in a high throughput manner (Houle et al. 2003;
123 Pitchers et al. 2013). Numerous studies have shown that mutational perturbations can have
124 profound effects on mean shape (Weber et al. 2005, Breuker et al. 2006; Debat et al. 2006;
125 Dworkin and Gibson 2006; Debat et al. 2009, Debat et al. 2011). The influence of mutational
126 perturbations on variance has been more mixed, with some studies seeing no effect, some find a
127 general increase (i.e., de-canalization; Debat et al. 2009, Debat et al. 2011), and others finding
128 both increase and decrease (Debat et al. 2006; Dworkin and Gibson 2006). Some evidence for the
129 effect of mutational and environmental perturbations on matrix orientation has been shown in
130 these studies as well. However, changes to covariance and shape have not been compared
131 directly, and other aspects of the covariance structure, such as eccentricity, have not been
132 included before.

133 To better understand the evolutionary lability of the **E** matrix, we used a set of induced
134 mutations measured against a common co-isogenic wild type (Dworkin and Gibson 2006; Debat
135 et al. 2009), as well as a panel of strains derived from natural populations, and examined the
136 relationship between mean shape and different aspects of the covariance. Thus, all strains were
137 reared under the same carefully-controlled conditions, and all mutant lines had the same genetic
138 background, minimizing unknown sources of variation. We observe that mean shape varies more
139 freely than the covariance structure, and that matrix orientation varies more than – and
140 independently from – its eccentricity and total variance.

141

142 Material and Methods

143 **Drosophila strains**

144 Insertional mutations (caused by the insertion of P-elements, marked with w^+) in genes
145 involved with the TGF-β and EGFR signaling pathways were provided from the Bloomington Stock
146 Center (Table 1). These represent a subset of the alleles that were first described in Dworkin and
147 Gibson (2006) and were chosen because of their important roles in the growth, patterning, and
148 shape determination of the *Drosophila* wing. All insertions were initially introgressed into the
149 wild-type lab strain Samarkand (Sam) for at least 10 generations. The Samarkand genotype was
150 marked with w^- (white eyes) so flies with insertions could be distinguished by a rescue of the eye
151 color phenotype. Introgressions were performed by repeated backcrossing of females bearing
152 the insertion to males of Sam. Prior to generating flies for the experiments described in this
153 study, the alleles were backcrossed to Sam for an additional four generations to remove any *de*
154 *novo* mutations that had accumulated in those lines (relative to Sam) since the original
155 introgression procedure. Selection was based entirely on the presence of the eye color marker,
156 precluding unwitting selection for wing phenotypes. All crosses were performed using standard
157 cornmeal-molasses media, in a 24° C Percival incubator on a 12/12-hr light/dark cycle with 60%
158 relative humidity (Dworkin and Gibson 2006). The mutant strains used here are part of a larger
159 unpublished study (Dworkin), and were chosen in part based on sample size (>50/genotype).

160 **Experimental setup**

161 After the additional four generations of backcrossing back to Samarkand, crosses between
162 each mutation and Samarkand were set up in vials, allowing females to lay for 2-3 days, so that

163 egg density was low (generally less than 60 flies/vial). The temperature of the incubator was
164 maintained at 24° C, and monitored carefully for fluctuations, and vial position was randomized
165 within the incubator on a daily basis to reduce any possible edge effects. As larvae crawled out of
166 the media, a piece of paper towel was added to each vial to provide additional pupation space.
167 After eclosion and sclerotization flies were separated, on the basis of eye color, into individuals
168 without the P-element-induced mutations (w-) and heterozygotes for the P-element, and then
169 stored in 70% ethanol.

170 ***Strains derived from natural populations from North Carolina and Maine***

171 In order to broaden the comparative basis we included a set of naturally-derived Iso-
172 female lines of *D. melanogaster*. Since too little is known about the scale of differences we might
173 expect for matrix properties, even for fruit flies, it helps to look at the range of differences found
174 among natural populations. Since shape has no natural zero (it is on interval scale), we chose Sam
175 as the point of comparison for all strains, mutants as well as naturally-derived.

176 The naturally-derived lines were established separately in the summer of 2004 at a Peach
177 orchard in West End, North Carolina (NC2) and in a blueberry field in Cherryfield, Maine (courtesy
178 of Marty Kreitman). To generate inbred lines, full-sib mating was performed for 15-20
179 generations (Goering et al 2009; Reed et al 2010). Flies were reared at 25C (+/- 1°C) in a 12/12
180 light/dark cycle at 50% humidity. We note that a different incubator was used for these lines
181 (compared to the mutant lines), but preliminary analyses for several of the P-element mutant
182 lines demonstrated highly consistent results with respect to mean shape (not shown). Seven
183 Maine and five NC2 lines, each with at least 51 males, were used for this study.

184 **Data collection**

185 A single wing from each fly was dissected and mounted in 70% glycerol (~20 wings per
186 replicate vial on average, >50 wings per genotype, males only; see Table 1). Wings were imaged
187 at 40X magnification on an Olympus DP30BW camera mounted on an Olympus BX51 microscope
188 using 'DP controller' V3.1.1 software. All images were saved as greyscale tiff files. To extract
189 landmark and semi-landmark data, we followed a modified protocol for the use of the
190 WINGMACHINE software (Houle et al. 2003) as detailed in Pitchers et al. (2013). First, we used
191 "tpsDig2" (Rohlf 2010) software to record the coordinates of the two starting landmarks needed
192 by WINGMACHINE. We then used WINGMACHINE to automatically fit nine B-splines to the veins
193 and margins of the wing (Fig. 2A). We reviewed each splined image and manually adjusted the
194 control points as necessary. The x and y coordinates of 14 landmarks and 34 semi-landmarks (Fig.
195 2A) were extracted using additional software (CPR) developed by Marquez and Houle (2014). The
196 data were checked for visual outliers on scatter plots, and putative outliers were examined, and
197 either fixed in WINGMACHINE or deleted. The resulting dataset was then passed on to **R** for
198 further analysis.

199 **Preparation of data**

200 All specimens from all genotypes were superimposed together using Generalized
201 Procrustes Analysis and projected into tangent space (function *gpagen* in the **R** package
202 *geomorph 2.1.1;* Adams and Otarola-Castillo 2013). The semi-landmarks were optimized using
203 minimum Procrustes distance. This resulted in a space of 96 Procrustes coordinates with 58
204 degrees of freedom (Zelditch et al. 2004). Replicate effect and the allometric effect of size were
205 removed from the shape coordinates by fitting a linear model, excluding their interaction, carried

206 out separately for each genotype. The predicted mean configuration was then added back to the
207 residuals to maintain the differences among genotypic means.
208 After regression, the dataset was passed on to the LORY program for calculating the
209 interpolated Jacobian-based data (Marquez et al. 2012). LORY uses spatial interpolation to
210 evaluate shape deformation at predetermined evaluation points throughout the wing. It starts by
211 creating a tessellated grid, with the landmarks and semi-landmarks as the vertices (Fig. 3A). The
212 centroids of the resulting triangles are considered the evaluation points. Given an interpolation
213 function, it then calculates the Jacobian matrix that describes the shape deformation at each
214 point (e.g., Fig. 3B). The log-transformed determinant of the Jacobian matrix quantifies the
215 amount of expansion or contraction at each evaluation point relative to the mean configuration.
216 Thus, the interpolation takes into account information from the whole configuration of
217 landmarks, as well as from the rest of the sample through the mean configuration, while
218 quantifying local shape changes. The Procrustes coordinates are mathematically transformed
219 from a single multidimensional trait into a multivariate set of traits. Just like interlandmark
220 distances (and unlike the Procrustes shape variables), each of the LORY variables can be
221 interpreted independently. The whole dataset can thus be analyzed using conventional
222 multivariate techniques (e.g., PCA, multivariate regression), yielding results that are more easily
223 interpretable as well (Marquez et al. 2012).
224 The interpolation function allows the researchers to incorporate and test an explicit
225 model for the distribution of shape changes across the configurations, rather than relying solely
226 on the Procrustes superimposition. The covariance matrix from the Procrustes residuals is
227 modified based on that explicit model. The most common interpolation function used in

228 geometric morphometrics studies is the Thin Plate Spline (TPS), which assumes the object is rigid
229 and therefore penalizes local deformations more than global ones. In this study we compared
230 TPS with another interpolation function implemented in LORY, the Elastic Body Spline (EBS),
231 which assumes the object is elastic and penalizes local deformations less than global ones
232 (Marquez et al. 2012). Therefore, like the Procrustes superimposition, TPS tends to spread the
233 variation more globally than EBS. Because TPS and EBS make different assumptions about the
234 distribution of variance among landmarks, they also make different assumptions about the
235 integration pattern of the wing. EBS largely implies less integration and more
236 compartmentalization. For this study system, we consider EBS to be a more suitable model *a*
237 *priori*, because the development of the fruit fly wing has been shown by most studies to proceed
238 in a relatively compartmentalized manner in which variation tends to be locally contained (Zecca
239 and Struhl 2002; Barrio and de Celis 2004; Cook et al. 2004; Martín and Morata 2006; Ziv et al.
240 2012; cf. Klingenberg and Zaklan 2002 and Debat et al 2003).
241 Each of the three dataset – the Procrustes coordinates and the LORY variables based on
242 TPS and EBS – was reduced to its first 30 dimensions using Principal Component Analysis,
243 covering 99% of the variation in the data (Fig. S1). All subsequent analyses, including estimation
244 of the covariance matrices, are based on these PCA scores. Thus, we have three sets for each of
245 these measures: mean shape distances, total variance, covariance distances (i.e., differences in
246 orientation), and eccentricity. However, results based on Procrustes data were highly correlated
247 with the Jacobian-based data, using both TPS and EBS (Figs. S2-S3). Therefore, we chose to
248 present below only results based on EBS, and provide results for the other two datasets in the
249 supplemental material. Thus, for *Drosophila* wing shape, using the LORY variables did not make a

substantial difference (compared to simply using Procrustes residuals). However, this could change for other cases, depending on the specific covariance structure of any given set of configurations (see examples in Marquez et al. 2012).

Quantification of variation and covariation

Mean shape distances were calculated as the Euclidean distance for both the interpolated data and the Procrustes-based data. Since the Procrustes data were projected to tangent space after superimposition (see above), this is equivalent to Procrustes distance. Variance-covariance (VCV) matrices were quantified based on their total variance, shape, and orientation (see Fig. 1). Total variance is the trace of the matrix (i.e., sum of its eigenvalues). Matrix shape was characterized based on its eccentricity (Jones et al. 2003), and calculated as the ratio of the largest eigenvalue to the total variance, which is the inverse of Kirkpatrick (2009)'s effective number of dimensions (Kirkpatrick 2009):

$$\text{Eccentricity} = \frac{\lambda_1}{\sum_{i=1}^{p} \lambda_i}$$

Where λ_1 is the largest eigenvalue, λ_i are the remaining i^{th} eigenvalues, and *p* is the number of variables in the VCV matrix. In addition, we calculated eccentricity as the relative standard deviation of the eigenvalues (rSDE), also known as integration level (Van Valen 1974; Pavlicev et al. 2009a; Haber 2011). As developed originally by Van Valen (1974) for covariance matrices, rSDE is also scaled by the total variance, thus measuring matrix shape only (Haber 2011, 2016). However, rSDE yielded essentially the same results as the above measure of eccentricity (Figs. S4 and S13), and will not be considered here further.

Similarity in matrix orientation was quantified using Random Skewers (Cheverud et al.

1983; Cheverud and Marroig 2007; Marroig et al. 2011), which measures the average similarity between two matrices in their response to random unit vectors. We used 5000 random vectors drawn from a normal distribution with mean 0 and variance 1, normed to unit length (Marroig et al. 2011). A distance metric was then calculated as $\sqrt{(1-r^2)}$, where r is the Random Skewers value for that pair of matrices, and normalized using Fisher's z-transformation (Jamniczky and Hallgrímsson 2009). In accordance with previous studies, this metric is referred to here as "covariance distance". Two other distance metrics were calculated as well: a modification of Krzanowski's (1979) metric following Zelditch et al. (2006), and the relative eigenvalues metric of Mitteroecker and Bookstein (2009). However, these two methods proved unstable for our dataset under resampling, yielding unreliable confidence intervals. At the same time, their estimated values followed the same pattern as Random Skewers (see online supplementary material). Moreover, Random Skewers is the only method of the three that is related to evolutionary theory (Hansen and Houle 2008). Therefore, we present below only results based on Random Skewers.

Confidence intervals for shape distances, total variance, and eccentricity, were estimated using a non-parametric bootstrap procedure with a BCa correction (DiCiccio and Efron 1996; Carpenter and Bithell 2000) using 999 iterations. The BCa correction was necessary because the pseudovalue distribution is expected to be biased upward when the statistic is bounded by zero, and to depend on its mean when bounded by both zero and one. These confidence intervals were used for evaluating significant differences as well. Confidence intervals for covariance distance were estimated using a Jackknife procedure. Each Jackknife pseudovalue was calculated by leaving out one specimen in one of the two samples that are being compared. The confidence

interval was calculated as the 95% of the Jackknife distribution, without a bias correction. The Jackknife was preferred in this case because the bootstrap (both parametric and non-parametric) consistently resulted in distributions that were highly biased upward, often excluding the observed value, invalidating the BCa correction. This is a commonly found phenomenon in similar studies and has not yet been properly addressed in the literature to the best of our knowledge. Here we found that Jackknife without a bias correction provided fairly symmetric distributions around the observed values with reasonably wide confidence intervals, whereas the bias-corrected jackknife yielded extremely narrow intervals that often excluded their observed value (i.e., they were "over-corrected").

In order to further compare covariance changes with shape changes we used Principal Coordinates Analysis to generate a shape space and a covariance space, based on all pairwise distances. In each of these spaces, genotypes are located based on how different they are from each other in either shape (i.e., shape space) or matrix orientation (i.e., covariance space). The two spaces where then superimposed together using a symmetric Procrustes superimposition including scaling (function *protest* in **R** package *vegan 2.0-10*; Oksansen et al. 2013). The superimposition ensures that the two spaces are comparable in terms of both magnitude and direction of change, just as it does with landmark configurations (Peres-Neto and Jackson 2001). The sum of squared deviations between the two spaces after superimposition provides a measure of correlation between them. In addition, we calculated the disparity of mutants and naturally-derived strains - for shape and covariance - as the sum of variances of their respective scores in the joint space, including all 20 PCoA axes. Although the configurations are scaled by their respective centroid size during the superimposition, the protest function then rescales the

315 rotated configuration (the covariance space in this case) proportionally to the target
316 configuration (the shape space in this case) so that their size (total variance) in the joint space is
317 comparable but not necessarily the same. This allows us to evaluate the disparity of one space
318 relative to the other.

319 We used MANOVA (Pillai's Λ) to test the effect of the developmental pathway on the
320 mutants' PCoA scores. As mentioned above, each genotype involved one mutation that targeted
321 either the TGF-β or the EGFR signaling pathway (see Table 1). This analysis allowed us to test
322 whether mutations clustered based on the pathway they targeted.

323 In order to estimate the effect of the high dimensionality of our data on sampling
324 variance, we repeated all analyses using a reduced dataset, for which we kept only the 12
325 landmarks and omitted all semi-landmarks. This dataset had 20 degree of freedom (12*2-4),
326 which is less than half our smallest sample size (45). This reduced dataset yielded very similar
327 results as the full dataset, for all properties, in terms of both the observed values and the
328 confidence intervals (Figs. S5-S7). The fact that the confidence intervals are similar is especially
329 relevant here, indicating directly that sampling variance is not affected by reducing the dataset
330 dimensionality from 56 to 20.

331 All analyses were carried out in **R** 3.2.4. All scripts and data will be made available on
332 DRYAD and on the Dworkin lab github repository.

333

334 Results

335 ***Both mean shape and matrix orientation vary considerably among mutant and***
336 ***naturally-derived strains.***

337 All mutants differ significantly in their mean shape from the Sam wild type (Table 2 and
338 Fig. 4; see also Fig S8), as determined by the lack of overlap between their confidence intervals
339 and the benchmark of zero distance from Sam (horizontal dashed line in Fig. 4). Most mutant
340 strains are more similar to the Sam wild type than any of the naturally-derived strains. Yet, two of
341 the mutants (genotypes 3045 and 10413, genes *Omb*^{md653} and Bs^{k07909} respectively) show
342 magnitudes of shape change as extreme as any of the naturally derived strains, suggesting that
343 the range of mutations included here likely provide a reasonable representation of what we
344 might observe in nature. The orientation of the covariance matrix of all mutant strains also differ
345 significantly from Sam (Table 2 and Fig. 5): none of the covariance distances include zero in their
346 confidence interval. Unlike mean shape, however, most of the mutant strains differ from Sam
347 just as much as the naturally-derived strains do.

348 ***Several mutant strains have lower environmental variance than their wild type***

349 Most mutations resulted in a lower total variance than their co-isogenic Sam wild-type
350 (Table 2 and Fig. 4). Six of the 12 mutant strains exclude the estimate for Sam from their
351 confidence interval (Table 2), thus indicating a significant difference. Two of the significant strains
352 (*Ptc*^{k02507}, *tkv*^{kg01923}) are part of the TGF-β pathway, and the other four (*Gap1*$^{mip\text{-}w[+]}$, *S*^{k09530},
353 *rho*^{kg07115}, *spi*^{s3547}) contribute to EGFR signaling (Table 1). The only strain that has substantially
354 and significantly higher total variance is *Omb*^{md653}, which also differs greatly in its mean shape. In
355 contrast, the total variance within each of the naturally-derived strains is mostly higher than the
356 total variance within Sam. In addition, there is a positive linear relationship between the total
357 variance within the mutant strains and their shape distance from Sam (Fig. 4).

358 ***A positive non-linear association between changes in mean shape and matrix***
359 ***orientation***.

360 There is a weak but positive association between covariance distance (i.e., difference in
361 matrix orientation) and shape distance (Fig. 5), relative to Sam. This relationship, however, is not
362 linear (Table 3), suggesting that the covariance distance is somewhat bounded at the upper
363 range. These findings are further supported by comparing the shape and covariance PCoA spaces
364 (Fig. 6, see also Figs S9-S12). These spaces are based on all pairwise comparisons, rather than
365 comparing to Sam only, and were superimposed for comparability. The superimposition scales
366 the two spaces to a common scale, centers, and rotates them so that they are comparable in
367 terms of both magnitude and direction of change. The grey solid lines indicate the deviation
368 between the two spaces. The sum of squared deviations between the two spaces is 0.669, and
369 the coefficient of determination is small yet significantly different from zero ($R^2 = 0.33$; $p \leq 0.001$
370 with 999 permutations). Thus, the two spaces are substantially different from each other, but not
371 unrelated. The first two PCoA axes cover 58% of the variation in the joint space (Fig. 6), while the
372 third and forth axes cover only 10% and 6% of the variation, respectively. The first PCo1 axis
373 mostly separates the naturally-derived strains from the mutants and Sam, for both shape and
374 covariance. The disparity of shape is more than twice as large as that of covariance (i.e.,
375 orientation differences) (0.042 and 0.020 respectively; calculated over all 20 PCoA dimensions
376 that resulted in eigenvalues larger than zero). The disparity of both shape and covariance is larger
377 for the naturally-derived strains than for the mutants (Fig.6A). The mutants do not cluster by
378 pathway (Fig. 6B), for either shape (Pillai's $\Lambda=0.44$, $p=0.24$) or covariance distance (Pillai's $\Lambda=0.34$,
379 $p=0.52$).

380 ***Most mutations do not influence matrix eccentricity, despite effects on mean shape,***
381 ***total variance, and matrix orientation***.
382 Most of the mutations do not have a substantial effect on the eccentricity of covariation
383 (i.e., matrix shape), and include the estimate for Sam well within their confidence interval (Table
384 2). In addition, there is no clear relationship between eccentricity and total variance (Fig. 7A). The
385 only mutations that caused a significant change in eccentricity are *aos*^{W11}, which increases
386 relative to Sam, and *sbb*^{BG01610}, which decreases relative to Sam (genotypes 2513 and 12772,
387 respectively). These mutations do not differ much from Sam for total variance and mean shape
388 (Fig. 7 and Table 2). *Omb*^{md653}, on the other hand, differs significantly and substantially from Sam
389 in its total variance and shape (see Table 2), and yet is very similar in its eccentricity to Sam. Most
390 of the naturally-derived strains have a higher eccentricity than Sam (Fig. 7A), showing a greater
391 difference from Sam than the mutants do. Similarly, there is no clear association between
392 eccentricity and shape distance from Sam (figure 7B), or between eccentricity and covariance
393 distance (i.e., orientation differences) from Sam (Fig. 7C). The same picture emerges when
394 eccentricity is calculated using the relative standard deviation of the eigenvalues (rSDE; Table 2
395 and Fig. S13), and based on Procrustes and TPS data (Figs. S14-S15).
396
397 Discussion
398 The pattern and magnitude of phenotypic covariation remains central to selection theory
399 (Robertson 1966; Price 1970; Lande and Arnold 1983). Although it cannot usually facilitate
400 adaptation by itself, the environmental component of phenotypic covariation – **E** – can impede
401 and divert the population response to selection (Burger 1986). Despite the importance of this for

402 understanding evolution, factors that influence the lability of **E** are poorly understood. For
403 univariate measures, environmental and genetic stressors have been shown to increase variance
404 in **E** in numerous situations (Dworkin 2005a; Hallgrímsson et al. 2009; Paaby and Rockman 2014;
405 and references therein), but little is known about how such perturbations influence other
406 properties of integration and covariation.

407 Using the wing of *Drosophila melanogaster* as a model system, we found that mean shape
408 and matrix orientation vary substantially among the co-isogenic mutants, as well as the naturally-
409 derived strains (Table 1 and Fig. 5). By superimposing the covariance space onto shape space we
410 were able to compare their distribution in terms of both magnitude and direction of change and
411 show that shape has changed to a greater extent than matrix orientation (Fig. 6). In contrast,
412 mutations had little influence on the total variance and eccentricity of the matrix. Whereas total
413 variance is positively associated with mean shape, it is not associated with eccentricity, nor is
414 eccentricity associated with matrix orientation. Together, these findings suggest that the
415 potential of the covariance structure to change is more limited than that of mean shape, and that
416 different properties of the covariance structure can change independently from each other. Thus,
417 even though these mutations greatly affect the location of genotypes in morphospace (i.e., mean
418 shape), the availability of different trait combinations to natural selection (matrix orientation) is
419 not affected as much. Moreover, the total amount of variation (matrix size) and its distribution
420 among the different trait combinations (eccentricity) are not affected much either. It is worth
421 noting that some of the environmental effects (such as rearing temperature and density) were
422 highly controlled in this study, both within and among strains. Thus, our experimental design may

423 represent a lower bound for the environmental contributions to covariation, and the patterns
424 observed are likely an underestimate of the possible effect sizes.
425 The common expectation is that mutational perturbations would cause a disruption to
426 the developmental system, leading to a higher variance and greater shape changes (Dworkin
427 2005b; Hall et al 2007; Hallgrímsson et al. 2009; Hill and Mulder 2010; and references therein).
428 Our findings are consistent with the common expectation in that greater decanalization is indeed
429 associated with larger changes to mean shape, but surprising in the sense that most mutants
430 have a lower variance than their co-isogenic wild type (albeit significant for only 6/12 genotypes),
431 and thus seem to be more canalized. If we had used linear measurements, we would have
432 expected to find a positive correlation between the mean and the variance merely due to scale.
433 However, such a technical relationship is not expected for Procrustes shape data because all
434 specimens are scaled by their centroid size during the superimposition. Thus, the positive
435 association we find likely reflects more than a technical relationship. In addition, the difference in
436 variance between strains is not likely to be due to differential survival, because the individuals
437 within each strain are genetically identical and viability was high enough for all genotypes. Thus,
438 the decrease in total variance within the mutants likely reflects a higher canalization of wing
439 shape. Dworkin and Gibson (2006) have also observed such mixed results for mutant strains,
440 using a smaller set of landmarks. A related study (Debat et al 2009) showed a more consistent
441 increase in total variance, even for some of the same mutations. However, the genetic
442 background of the strains in Debat et. al (2009) was Oregon-R (rather than Samarkand), which
443 has been shown to be more sensitive to mutational perturbations under some conditions (Chari
444 & Dworkin 2013).

445 Increases in the environmental component of phenotypic variance have been explained in
446 previous studies by a nonlinear relationship between the trait mean and its underlying
447 developmental parameters (Klingenberg and Nijhout 1999; Hallgrímsson et al. 2009). The
448 nonlinearity arises from the expectation that the variation expressed around the mean
449 phenotype can vary for different phenotypes, reflecting different levels of canalization. A steeper
450 slope around a given phenotype implies a greater amount of variation, and therefore less
451 canalization, associated with that phenotype. It is often assumed that when development is
452 perturbed the mean would shift to a less canalized section on the curve and express greater
453 variation. It is possible, however, that the mean would shift into more canalized regions, rather
454 than less canalized, resulting in a decrease of variance rather than an increase (Hill and Mulder
455 2010). In other words, Sam is more sensitive to micro-environmental influences than the mutants
456 and less sensitive than the natural strains, at least in our study. It is also possible to have system-
457 wide canalizing factors (such as chaperon proteins; Rutherford and Lindquist 1998) operating
458 simultaneously, masking the effects of local perturbations and stabilizing the variance
459 (Hallgrímsson et al. 2006). Or, the mutations could increase redundancy and complexity thus
460 increasing stability (Hallgrímsson et al. 2006; Levy and Siegal 2008). All of these possible
461 explanations would be consistent also with the positive association we found between mean
462 shape changes and total variance.
463 A possible explanation for the higher variance within the natural strains (relative to Sam)
464 could be related to the time and amount of inbreeding among the strains (Whitlock & Fowler
465 1999; Whitlock et al. 2002). Inbreeding in *Drosophila* does represent a form of "genetic stress" as
466 deleterious recessive alleles are made homozygous (and their effects are phenotypically

467 expressed), often leading to an increase in environmental variance. We would therefore expect
468 the natural strains to be de-canalized as well. However, the Sam wild-type has also undergone
469 extensive inbreeding, and is near isogenic (Chandler et al. 2014). Thus, a satisfactory explanation
470 for this particular observation eludes us.

471 In accord with previous studies (Dworkin and Gibson 2006; Debat et al. 2009; Jamniczky
472 and Hallgrímsson 2009), covariance matrices of the mutant strains do not cluster according to the
473 developmental pathway that their underlying mutations presumably affect. This could point to
474 the complexity of the developmental system, further supporting the palimpsest model suggested
475 by Hallgrímsson et al. (2009). It could also mean that these pathways do not contribute to the
476 structure of **E** as effectively and consistently as previously postulated. A detailed analysis of the
477 covariance pattern is needed to further investigate the extent to which the modularity structure
478 of the fruit fly wing follows *a priori* developmental models, and how it is disrupted by
479 developmental perturbations.

480 It is difficult to compare differences in eccentricity to differences in orientation directly
481 because of their different dimensionality and somewhat different method for calculating the
482 confidence intervals. However, mutant eccentricity varies from that of Sam within a smaller
483 portion of its possible range (0.2-0.43; 23%) compared to how much their orientation varies from
484 Sam (0.37-0.77; 40% without the Fisher's *z*-transformation). These findings are in accord with
485 simulations by Jones et al. (2003, 2004, 2012) for **G**, suggesting that eccentricity is likely to vary
486 less than orientation for a given population size under most combinations of genetic
487 architecture, selection regimes, and phylogenetic scale. It is also consistent with empirical
488 evidence for **P** from Haber (2015a,b), showing that matrix orientation has varied greatly among

489 closely-related ruminant species, whereas eccentricity has remained largely the same throughout
490 most of bovid history, and has only varied within 33% of its possible range among other
491 ruminants.

492 As with many studies of covariation, sampling variance could have a substantial impact on
493 our results, considering the high dimensionality of the data, and especially on estimates of
494 covariance. Several studies (i.e., Hill and Thompson 1978; Meyer and Kirkpatric 2008; Pavlicev et
495 all. 2009b) have shown that the leading eigenvector tends to be overestimated due to sampling
496 variance, and the trailing eigenvalues underestimated. This would affect eccentricity the most.
497 Simulations carried by Haber (2011) indicate that eccentricity is only biased (overestimated) for
498 low values (rSDE<0.2) with sample sizes lower than 45 and number of variables 35 or lower. Our
499 sample sizes are all higher than 45, and most are higher than 60, with 56 degrees of freedom, so
500 largely equivalent to those simulations, and all of our rSDE values are higher than 0.2. There is no
501 reason to expect other measures of eccentricity to be affected differently as they are all very
502 tightly correlated and reflect the same property (the distribution of the eigenvalues). In addition,
503 the power analysis in Haber (2011) indicates that increasing sample size above 40 adds little to
504 the statistical power of rSDE. Therefore, it is reasonable to assume that our eccentricity estimates
505 are largely unbiased by sampling variance. With regards to other properties, mean shape and
506 total variance are scalars, known to be relatively robust to sampling (Zelditch et al 2004). In
507 contrast, matrix orientation is probably the most sensitive property, but it is impossible to say
508 how sampling variance would affect it as it depends largely on the covariance structure itself.
509 However, since sample size is similar for all mutant lines, its effect should be about the same as
510 well. Moreover, repeating the analyses with a reduced dataset of 12 landmarks (see methods),

511 yielded very similar results as the full dataset for all properties, in terms of both the point
512 estimates and their confidence intervals (Figs. S5-S7). The fact that the confidence intervals are
513 similar is informative in this context as it indicates that the sampling variance is not substantially
514 affected by reducing the dataset from 56 to 20. Pitchers et al. (2014) also found that the number
515 of families (the relevant sample size for that study) had relatively modest effect on various
516 measures of matrix properties of **G**.

517 To conclude, previous studies have shown an increase in variance to be clearly associated
518 with decanalization. In this study, however, we find that mutants altered the orientation of the
519 covariance matrix more often than its total variance. Mutations do not only affect trait means
520 and variances, but aspects of covariances as well, though this has not been part of the general
521 formulation so far. Thus, our study suggests that it might be useful to consider a more general
522 concept of decanalization.

523

524 **Acknowledgements**: We thank Ellen Larsen, Jeff Conner, Will Pitchers, Vincent Debat,
525 Adrien Perrard, Handling Editor Ruth Shaw, Associate Editor Ricardo Azevedo, and two
526 anonymous reviewers for comments that improved clarity of this manuscript. This material is
527 based in part on work supported by the National Science Foundation under Cooperative
528 Agreement No. DBI-0939454. Any opinions, findings and conclusions or recommendations
529 expressed in this material are those of the author(s) and do not necessarily reflect the views of
530 the National Science Foundation. This work was supported by NIH grant 1R01GM094423-01 to
531 ID. Funding for AH was provided by the Council of Higher Education, Israel, and BEACON Center
532 for the Study of Evolution in Action at Michigan State University.

533

533 Literature cited

534 Adams, D. C., and E. Otarola-Castillo. 2013. geomorph: an R package for the collection and
535 analysis of geometric morphometric shape data. Methods in Ecology and Evolution 4:393-
536 399.

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

539 Barrio, R., and J. F. de Celis. 2004. Regulation of *spalt* expression in the *Drosophila* wing blade in
540 response to the Decapentaplegic signaling pathway. Proceedings of the National Academy
541 of Sciences of the United States of America 101:6021-6026.

542 Breuker, C. J., J. S. Patterson, and C. P. Klingenberg. 2006. A Single Basis for Developmental
543 Buffering of *Drosophila* Wing Shape. PLoS ONE 1:e7.

544 Burger, R. 1986. Constraints for the evolution of functionally coupled characters: A nonlinear
545 analysis of a phenotypic model. Evolution 40:182-193.

546 Carpenter, J., and J. Bithell. 2000. Bootstrap confidence intervals: when, which, what? A
547 practical guide for medical statisticians. Statistics in Medicine 19:1141-1164.

548 Chari, S., and I. Dworkin. 2013. The Conditional Nature of Genetic Interactions: The
549 Consequences of Wild-Type Backgrounds on Mutational Interactions in a Genome-Wide
550 Modifier Screen. PLoS Genetics 9:e1003661.

551 Chandler, C. H., S. Chari, D. Tack, and I. Dworkin. 2014. Causes and Consequences of Genetic
552 Background Effects Illuminated by Integrative Genomic Analysis. Genetics 196:1321-1336.

553 Cheverud, J. M., J. J. Rutledge, and W. R. Atchley. 1983. Quantitative Genetics of Development:
554 Genetic Correlations Among Age-Specific Trait Values and the Evolution of Ontogeny.

555 Evolution 37:895-905.

556 Cheverud, J.M., and G. Marroig. 2007. Comparing covariance matrices: random skewers method
557 compared to common principal components model. Genetics and Molecular Biology 30:
558 461-469.

559 Cook, O., B. Biehs, and E. Bier. 2004. *brinker* and *optomotor-blind* act coordinately to initiate
560 development of the L5 wing vein primordium in *Drosophila*. Development 131:2113-2124.

561 de Oliveira, F. B., A. Porto, and G. Marroig. 2009. Covariance structure in the skull of Catarrhini:
562 a case of pattern stasis and magnitude evolution. Journal of Human Evolution 56:417-430.

563 Debat, V., S. Bloyer, F. Faradji, N. Gidaszewski, N. Navarro, P. Orozco-terWengel, V. Ribeiro, C.
564 Schlötterer, J. S. Deutsch, and F. Peronnet. 2011. Developmental Stability: A Major Role
565 for *Cyclin G* in *Drosophila melanogaster*. PLoS Genetics 7:e1002314.

566 Debat, V., A. Debelle, and I. Dworkin. 2009. Plasticity, canalization, and developmental stability
567 of the *Drosophila* wing: joint effects of mutations and developmental temperature.
568 Evolution 63:2864-2876.

569 Debat, V., C. C. Milton, S. Rutherford, C. P. Klingenberg, and A. A. Hoffmann. 2006. Hsp90 and
570 the quantitative variation of wing shape in *Drosophila melanogaster*. Evolution 60:2529-
571 2538.

572 Debat, V., M. Béagin, H. Legout, and J. R. David. 2003. Allometric and nonallometric components
573 of Drosophila wing shape respond differently to developmental temperature. Evolution
574 57(12):2773-2784

575 DiCiccio, T. J., and B. Efron. 1996. Bootstrap Confidence Intervals. Statistical Science 11:189-212.

576 Dworkin, I. 2005a. A study of canalization and developmental stability in the sternopleural

577 bristle system of *Drosophila melanogaster*. Evolution 59:1500-1509.

578 Dworkin, I. 2005b. Canalization, cryptic variation, and developmental buffering: a critical
579 examination and analytical perspective. Pp. 131-158 *in* B. Hallgrímsson and B. K. Hall, eds.
580 Variation, a central concept in biology

581 Dworkin, I., and G. Gibson. 2006. The EGF-R and TGF-ß signal transduction pathways contribute
582 to variation for wing shape in *Drosophila melanogaster*. Genetics 173:1417–1431.

583 Goering, L. M., P. K. Hunt, C. Heighington, C. Busick, P. Pennings, J. Hermisson, S. Kumar, and G.
584 Gibson. 2009. Association of orthodenticle with natural variation for early embryonic
585 patterning in *Drosophila melanogaster*. Journal of Experimental Zoology Part B: Molecular
586 and Developmental Evolution 312B:841-854.

587 Haber, A. 2011. A Comparative Analysis of Integration Indices. Evolutionary Biology 38:476-488.

588 Haber, A. 2015a. The Evolution of Morphological Integration in the Ruminant Skull. Evolutionary
589 Biology 42:99-114.

590 Haber, A. 2016. Phenotypic Covariation And Morphological Diversification In The Ruminant Skull.
591 American Naturalist 187(5):576-591.

592 Hall, M. C., I. Dworkin, M. C. Ungerer, and M. Purugganan. 2007. Genetics of
593 microenvironmental canalization in *Arabidopsis thaliana*. Proceedings of the National
594 Academy of Sciences 104:13717.

595 Hallgrímsson, B., J. J. Y. Brown, A. F. Ford-Hutchinson, H. D. Sheets, M. L. Zelditch, and F. R. Jirik.
596 2006. The brachymorph mouse and the developmental-genetic basis for canalization and
597 morphological integration. Evolution & Development 8(1):61-73.

598 Hallgrímsson, B., H. A. Jamniczky, N. Young, C. Rolian, T. Parsons, J. Boughner, and R. Marcucio.

599 2009. Deciphering the Palimpsest: Studying the Relationship Between Morphological
600 Integration and Phenotypic Covariation. Evolutionary Biology 36:355-376.
601 Hansen, T. F., and D. Houle. 2008. Measuring and comparing evolvability and constraint in
602 multivariate characters. Journal of Evolutionary Biology 21:1201-1219.
603 Hill, W. G., and H. A. Mulder. 2010. Genetic analysis of environmental variation. Genetic Research
604 92(5):381–395.
605 Hill, W. G., and R. Thompson. 1978. Probabilities of non-positive definite between-group or
606 genetic covariance matrices. Biometrics 34(3):429-439
607 Hohenlohe, P. A., and S. J. Arnold. 2008. MIPoD: A Hypothesis-Testing Framework for
608 Microevolutionary Inference from Patterns of Divergence. The American Naturalist
609 171:366-385.
610 van Homrigh, A., M. Higgie, K. Mcguigan, and M. W. Blows. 2007. The Depletion of Genetic
611 Variance by Sexual Selection. Current Biology 17(6): 528–532.
612 Houle, D., J. Mezey, P. Galpern, and A. Carter. 2003. Automated measurement of Drosophila
613 wings. BMC Evolutionary Biology 3:25.
614 Jamniczky, H. A., and B. Hallgrímsson. 2009. A Comparison of Covariance Structure in Wild and
615 Laboratory Muroid Crania. Evolution 63:1540-1556.
616 Jones, A. G., S. J. Arnold, and R. Bürger. 2003. Stability of the G-Matrix in a Population
617 Experiencing Pleiotropic Mutation, Stabilizing Selection, and Genetic Drift. Evolution
618 57:1747-1760.
619 Jones, A. G., S. J. Arnold, and R. Bürger. 2004. Evolution and Stability of the G-Matrix on a
620 Landscape with a Moving Optimum. Evolution 58:1639-1654.

621 Jones, A. G., S. J. Arnold, R. Bürger, and C. Goodnight. 2007. The mutation matrix and the
622 evolution of evolvability. Evolution 61:727-745.
623 Jones, A. G., R. Bürger, S. J. Arnold, P. A. Hohenlohe, and J. C. Uyeda. 2012. The effects of
624 stochastic and episodic movement of the optimum on the evolution of the G-matrix and
625 the response of the trait mean to selection. Journal of Evolutionary Biology 25:2210-2231.
626 Kirkpatrick, M. 2009. Patterns of quantitative genetic variation in multiple dimensions. Genetica
627 136:271-284.
628 Klingenberg, C. P., and H. F. Nijhout. 1999. Genetics of Fluctuating Asymmetry: A Developmental
629 Model of Developmental Instability. Evolution 53:358-375.
630 Klingenberg, C. P., and S. D. Zaklan. 2000. Morphological integration between developmental
631 compartments in the drosophila wing. Evolution 54(4):1273-1285
632 Krzanowski, W. J. 1979. Between-groups comparison of principal components. Journal of the
633 American Statistical Association 74(367):703-707.
634 Lande, R. 1979. Quantitative Genetic Analysis of Multivariate Evolution, Applied to Brain: Body
635 Size Allometry. Evolution 33:402-416.
636 Lande, R., and S. J. Arnold. 1983. The measurement of selection on correlated characters.
637 Evolution 37:1210-1226.
638 Levy, S. F., and M. L. Siegal. 2008. Network Hubs Buffer Environmental Variation in
639 *Saccharomyces cerevisiae*. PLoS Biology 6:e264.
640 Marquez, E.J.; Houle, D. 2014. CPR: A WINGMACHINE tool for data processing and visualization.
641 Florida State University, Tallahassee, FL. (http://bio.fsu.edu/~dhoule/wings.html)
642 Marquez, E. J., R. Cabeen, R. P. Woods, and D. Houle. 2012. The Measurement of Local Variation

in Shape. Evolutionary Biology 39:419-439.

Marroig, G., D. Melo, A. Porto, H. Sebastião, and G. Garcia. 2011. Selection Response Decomposition (SRD): A New Tool for Dissecting Differences and Similarities Between Matrices. Evolutionary Biology 38:225-241.

Martín, F. A., and G. Morata. 2006. Compartments and the control of growth in the Drosophila wing imaginal disc. Development 133:4421-4426.

Meyer, K., and M. Kirkpatrick. 2008. Perils of Parsimony: Properties of Reduced-Rank Estimates of Genetic Covariance Matrices. Genetics 180(2):1153–1166.

Mitteroecker, P., and F. Bookstein. 2009. The ontogenetic trajectory of the phenotypic covariance matrix, with examples from craniofacial shape in rats and humans. Evolution 63:727-737.

Oksanen, J., F. G. Blanchet, R. Kindt, P. Legendre, P. R. Minchin, R. G. O'Hara, G. L. Simpson, P. Solymos, H. M. H. Stevens, and H. Wagner. 2013. *vegan*: Community Ecology Package.

Paaby, A. B., and M. V. Rockman. 2014. Cryptic genetic variation: evolution's hidden substrate. Nature Reviews Genetics 15:247-258.

Pavlicev, M., J. Cheverud, and G. Wagner. 2009a. Measuring morphological integration using eigenvalue variance. Evolutionary Biology 36:157-170.

Pavlicev, M., G. Wagner, and J. Cheverud. 2009b. Measuring evolutionary constraints through the dimensionality of the phenotype: adjusted bootstrap method to estimate rank of phenotypic covariance matrices.

Peres-Neto, P.R., and D.A. Jackson. 2001. How well do multivariate data sets match? The advantages of a Procrustean superimposition approach over the Mantel test. Oecologia 129: 169-178

Pitchers, W., J. E. Pool, I. Dworkin: Altitudinal clinal variation in wing size and shape in African *Drosophila melanogaster*: one cline or many? Evolution 2013, 67:438–452

Pitchers, W., J. B. Wolf, T. Tregenza, J. Hunt, and I. Dworkin. (2014). Evolutionary rates for multivariate traits: the role of selection and genetic variation. Philosophical Transactions of the Royal Society B: Biological Sciences, 369(1649):20130252.

Porto, A., F. de Oliveira, L. Shirai, V. De Conto, and G. Marroig. 2009. The Evolution of Modularity in the Mammalian Skull I: Morphological Integration Patterns and Magnitudes. Evolutionary Biology 36:118-135.

Price, G. R. 1970. Selection and covariance. Nature 227:520-521.

R Development Core Team. 2016. R: A Language and Environment for Statistical Computing. http://www.R-project.org.

Reed, L. K., S. Williams, M. Springston, J. Brown, K. Freeman, C. E. DesRoches, M. B. Sokolowski, and G. Gibson. 2010. Genotype-by-Diet Interactions Drive Metabolic Phenotype Variation in *Drosophila melanogaster*. Genetics 185:1009-1019.

Revell, L. J. 2007. The G matrix under fluctuating correlational mutation and selection. Evolution 61:1857-1872.

Robertson, A. 1966. A mathematical model of the culling process in dairy cattle. Animal Science 8:95-108.

Rohlf, F. 2010. tpsDig, digitize landmarks and outlines. 2.16. Department of ecology and

685 evolution. Stony Brook: State University

686 Rutherford, S. L., and S. Lindquist. 1998. Hsp90 as a capacitor for morphological evolution.

687 Nature 396:336-342.

688 Van Valen, L. M. 1974. Multivariate structural statistics in natural history. Journal of Theoretical

689 Biology 45:235-247.

690 Waddington, C. H. 1952. Selection of the Genetic Basis for an Acquired Character. Nature

691 169:625-626.

692 Walsh B., and M. W. Blows. 2009. Abundant Genetic Variation plus Strong Selection =

693 Multivariate Genetic Constraints: A Geometric View of Adaptation. Annu. Rev. Ecol. Evol.

694 Syst, 40:41–59

695 Weber, K., N. Johnson, D. Champlin, and A. Patty. 2005. Many P-Element Insertions Affect Wing

696 Shape in *Drosophila melanogaster*. Genetics 169:1461-1475.

697 Whitlock, M. C., P. C. Phillips, and K. Fowler. 2002. Persistence of changes in the genetic

698 covariance matrix after a bottleneck. Evolution 56:1968-1975.

699 Whitlock, M. C., and K. Fowler. 1999. The Changes in Genetic and Environmental Variance With

700 Inbreeding in *Drosophila melanogaster*. Genetics 152:345-353.

701 Zecca, M., and G. Struhl. 2002. Subdivision of the *Drosophila* wing imaginal disc by EGFR-

702 mediated signaling. Development 129:1357-1368.

703 Zelditch, M. L., Swiderski, D. L., Sheets, H. D., and Fink, W. L. 2004 Geometric Morphometrics for

704 Biologists. Elsvier Academic Press.

705 Zelditch, M. L., J. Mezey, H. D. Sheets, B. L. Lundrigan, and G. J. Theodore. 2006. Developmental

706 regulation of skull morphology II: ontogenetic dynamics of covariance. Evolution &

707 Development 8:46-60.

708 Ziv, O., R. Finkelstein, Y. Suissa, T. Dinur, G. Deshpande, and O. Gerlitz. 2012. Inverse regulation

709 of target genes at the brink of the BMP morphogen activity gradient. Journal of Cell

710 Science 125:5811-5818.

711
712

712 **Figure 1:** Illustration of the different matrix properties. Ellipses represent 95% confidence interval
713 of a normal bivariate distribution, and arrows (ellipse axes) represent orthogonal axes of
714 covariaton. Matrices are presented as bivariate for illustration, but the principles are the same
715 for multivariate. Ellipses A, B, and C represent a change in only one property each, relative to the
716 ellipse in the middle. A, lower total variance; B, perpendicular orientation; C, higher eccentricity.
717 **Figure 2: A,** A representative wing image with fitted splines and the location of landmarks and
718 semilandmarks along them. **B,** Individual variation in the sample (grey) relative to Sam (black),
719 mutant means (orange), and means for naturally-derived lines (purple).
720 **Figure 3:** Illustrating the LORY method for interpolating shape changes from Procrustes data of
721 genotype Bs^{k07909}. **A**, the tessellation scheme for choosing evaluation points. The landmarks and
722 semi-landmarks (dots) are used as vertices. The centroids of the resulting triangles (crosses) are
723 the final evaluation points. **B**, shape differences between the genotype mean and Sam inferred
724 by calculating the Jacobians around each evaluation point, scaled by 100 (crosses and crossed
725 circles). Bars represent the direction and change of differences between Sam and the sample
726 mean at the landmarks and semi-landmarks based on the Procrustes data, scaled by 10X.
727 **Figure 4:** Mean-shape distance from Sam plotted against the total variance within each strain.
728 Bars are 95% confidence intervals. Dashed lines represent values for Sam. Labeled mutants are
729 mentioned in the text.
730 **Figure 5:** Mean-shape distance plotted against covariance distance (i.e, difference in matrix
731 orientation; estimated by Random Skewers), relative to Sam. Bars are 95% confidence intervals.
732 Labeled mutants are mentioned in the text.

733 **Figure 6:** The first two axes of the covariance space superimposed onto shape space. Both spaces
734 are based on all pairwise distances. Solid grey lines indicate the deviation between the two
735 spaces for a given genotype. **A,** All genotypes; **B,** Mutants only, including Sam. Mutants that
736 affect the TGF-β pathway are marked by *t*. All other mutants affect the EGFR pathway. nWT,
737 naturally-derived wild types; mt, mutants. Labeled mutants are mentioned in the text.
738 **Figure 7:** The association between eccentricity and (**A**) total variance within each genotype, (**B**)
739 their mean-shape distance from Sam, and (**C**) their covariance distance from Sam. Dashed line
740 represents the eccentricity of Sam. The apparent positive relationship between eccentricity and
741 total variance is no greater than expected by chance based on parametric simulations. Labeled
742 mutants are mentioned in the text.

Table 1: Information about the genotypes included in this study

Genotype	N	Wild Type	Allele	Gene		Pathway
Sam	175	Sam	Wild type	Wild type		-
10385	81	Sam	k05115	*Epidermal growth factor Receptor*	*Egfr*^{k05115}	Egfr
10413	88	Sam	k07909	*Blistered*	*Bs*^{k07909}	Egfr
10418	84	Sam	k09530	*Star*	*S*^{k09530}	Egfr
10462	89	Sam	s3547	*spitz*	*spi*^{s3547}	Egfr
10514	83	Sam	k02507	*Patched*	*Ptc*^{k02507}	Hh
12772	90	Sam	BG01610	*scribbler*	*sbb*^{BG01610}	TGF-β
14321	89	Sam	kg07115	*rhomboid*	*rho*^{kg07115}	Egfr
14403	76	Sam	kg01923	*thickveins*	*tkv*^{kg01923}	TGF-β
2513	86	Sam	W11	*argos*	*aos*^{W11}	Egfr
3045	92	Sam	md653	*Optometer blind/Bifid*	*Omb*^{md653}	TGF-β
6342	92	Sam	225-3	*crossveinless-2*	*cv-2*$^{225-3}$	TGF-β
6372	82	Sam	mip-w[+]	*GTPase activating protein1*	*Gap1*$^{mip-w[+]}$	Egfr
MA10	63	MA	WT	WT	-	-
MA12	62	MA	WT	WT	-	-
MA44	52	MA	WT	WT	-	-
MA53	55	MA	WT	WT	-	-
MA61	56	MA	WT	WT	-	-
MA70	57	MA	WT	WT	-	-
MA71	55	MA	WT	WT	-	-
NC2	54	NC	WT	WT	-	-
NC30	51	NC	WT	WT	-	-
NC53	58	NC	WT	WT	-	-
NC9	51	NC	WT	WT	-	-
NC94	61	NC	WT	WT	-	-

Table 2: Matrix properties of all genotypes, and their distance from Sam in terms of matrix orientation (cov.dis) and mean shape (shape.dis). All measures are based on Jacobians, using EBS, and are therefore proportional to the mean configuration of that genotype.

Genotype	**Total variance**	**rSDE**	**Eccentricity**[a]	**cov.dis**[b]	**shape.dis**
Sam (wild type)	0.13 (0.11,0.15)	0.27 (0.26,0.28)	0.21 (0.17,0.23)	0 (0,0)	0 (0,0)
aos^{W11}	0.12 (0.11,0.15)	*0.35 (0.28,0.40)	*0.34 (0.25,0.42)	*0.57 (0.56,0.58)	*0.25 (0.22,0.29)
Omb^{md653}	*0.17 (0.16,0.20)	0.28 (0.27,0.29)	0.22 (0.19,0.25)	*0.68 (0.67,0.69)	*1.23 (1.19,1.27)
cv-2$^{225-3}$	0.12 (0.10,0.14)	*0.29 (0.28,0.30)	0.22 (0.18,0.26)	*0.46 (0.45,0.47)	*0.22 (0.19,0.25)
Gap1$^{mip-w[+]}$	*0.11 (0.10,0.12)	0.29 (0.27,0.30)	0.24 (0.19,0.29)	*0.58 (0.57,0.59)	*0.42 (0.39,0.45)
Egfr^{k05115}	0.12 (0.11,0.14)	0.30 (0.27,0.31)	0.24 (0.19,0.27)	*0.51 (0.51,0.53)	*0.23 (0.20,0.25)
Bs^{k07909}	0.13 (0.12,0.14)	0.30 (0.27,0.31)	0.26 (0.20,0.31)	*0.57 (0.57,0.58)	*0.63 (0.60,0.65)
S^{k09530}	*0.10 (0.09,0.12)	0.29 (0.26,0.30)	0.22 (0.19,0.27)	*0.44 (0.43,0.44)	*0.13 (0.11,0.14)
spi^{s3547}	*0.10 (0.09,0.11)	0.27 (0.26,0.27)	0.21 (0.17,0.24)	*0.52 (0.51,0.52)	*0.38 (0.34,0.41)
Ptc^{k02507}	*0.11 (0.10,0.12)	0.28 (0.25,0.28)	0.22 (0.18,0.25)	*0.55 (0.54,0.56)	*0.54 (0.51,0.56)
sbb^{BG01610}	0.11 (0.10,0.13)	0.27 (0.27,0.27)	*0.19 (0.17,0.19)	*0.50 (0.50,0.51)	*0.37 (0.34,0.40)
rho^{kg07115}	*0.10 (0.09,0.11)	0.28 (0.27,0.28)	0.21 (0.17,0.22)	*0.39 (0.39,0.40)	*0.17 (0.14,0.19)
tkv^{kg01923}	*0. 10 (0.09,0.12)	*0.30 (0.28,0.30)	0.23 (0.19,0.26)	*0.61 (0.60,0.61)	*0.30 (0.27,0.33)
MA10	0.15 (0.13,0.18)	0.35 (0.31,0.38)	0.30 (0.23,0.40)	0.66 (0.65,0.67)	1.00 (0.96,1.04)
MA12	0.15 (0.13,0.16)	0.34 (0.31,0.36)	0.29 (0.22,0.36)	0.59 (0.58,0.60)	0.96 (0.92,0.99)
MA44	0.15 (0.13,0.18)	0.34 (0.30,0.36)	0.28 (0.21,0.36)	0.58 (0.57,0.59)	0.91 (0.87,0.94)
MA53	0.17 (0.15,0.20)	0.33 (0.31,0.34)	0.28 (0.21,0.34)	0.62 (0.61,0.63)	1.04 (0.99,1.08)
MA61	0.13 (0.12,0.15)	0.35 (0.32,0.36)	0.29 (0.23,0.34)	0.69 (0.67,0.69)	1.15 (1.11,1.18)
MA70	0.12 (0.10,0.14)	0.30 (0.28,0.30)	0.22 (0.18,0.23)	0.71 (0.70,0.72)	1.11 (1.08,1.14)
MA71	0.17 (0.15,0.21)	0.30 (0.29,0.30)	0.22 (0.19,0.22)	0.69 (0.68,0.70)	1.41 (1.36,1.45)
NC2	0.12 (0.11,0.13)	0.28 (0.28,0.28)	0.20 (0.17,0.21)	0.82 (0.81,0.83)	1.15 (1.12,1.19)
NC30	0.12 (0.10,0.15)	0.29 (0.29,0.29)	0.21 (0.18,0.22)	0.72 (0.71,0.73)	0.89 (0.86,0.92)
NC53	0.13 (0.11,0.17)	0.31 (0.28,0.31)	0.22 (0.18,0.23)	0.81 (0.80,0.82)	1.06 (1.03,1.09)
NC9	0.15 (0.13,0.19)	0.39 (0.32,0.48)	0.37 (0.28,0.51)	0.73 (0.72,0.75)	0.54 (0.49,0.58)
NC94	0.21 (0.19,0.24)	0.44 (0.37,0.48)	0.42 (0.34,0.48)	0.98 (0.97,0.99)	1.02 (0.97,1.07)

* Confidence interval excludes the estimate for Sam; for mutant strains only

[a] Measured as the ratio between the largest eigenvalue and the total variance.

[b] Difference in orientation based on Random Skewers.

Table 3: Non-linear regression of shape distances on covariance distances (relative to Sam)

	Estimates	t value	P value
Intercept	-2.83 (-4.75, -0.92)	-3.07	0.0057 **
Linear term	9.07 (3.20, 14.94)	3.21	0.0042 **
Non-linear term	-5.19 (-9.57, -0.81)	-2.47	0.022 *

Model: $Y = X+X^2$; Y is shape distance; X is covariance distance

Tested sequentially so that the non-linear term is added to the linear term

Residual standard error: 0.25 on 21 degrees of freedom

Multiple R-squared: 0.63, Adjusted R-squared: 0.60

F-statistic: 18.29 on 2 and 21 DF, p-value: 2.52e-05

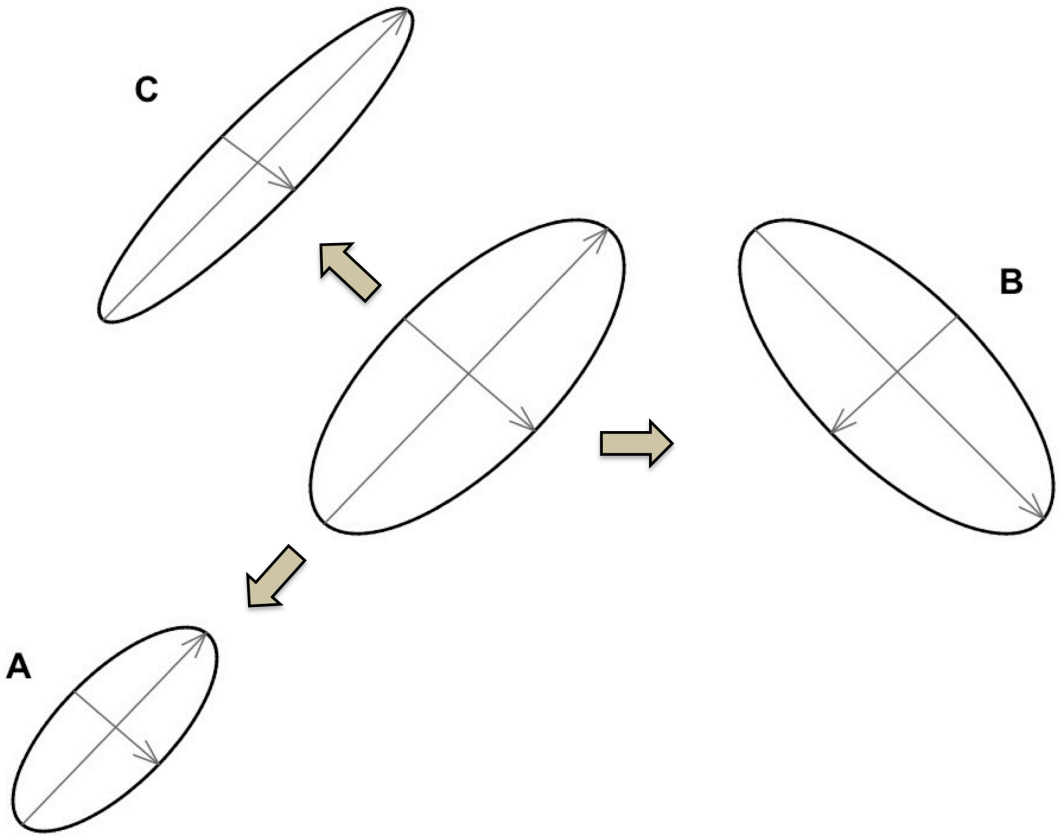
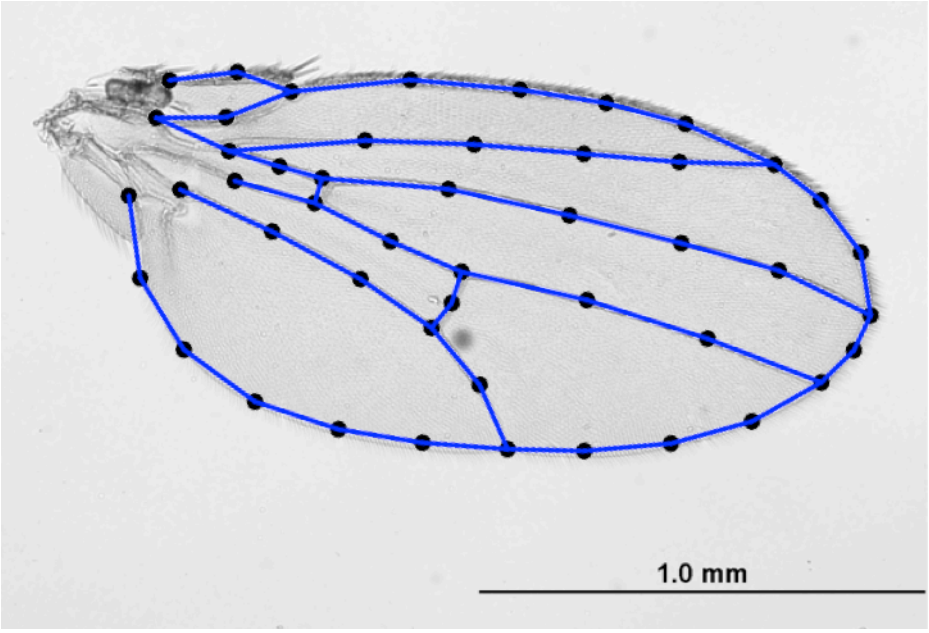


Figure 1: Illustration of the different matrix properties. Ellipses represent 95% confidence interval of a normal bivariate distribution, and arrows (ellipse axes) represent orthogonal axes of covariaton. Matrices are presented as bivariate for illustration, but the principles are the same for multivariate. Ellipses A, B, and C represent a change in only one property each, relative to the ellipse in the middle. A, lower total variance; B, perpendicular orientation; C, higher eccentricity.

A



B

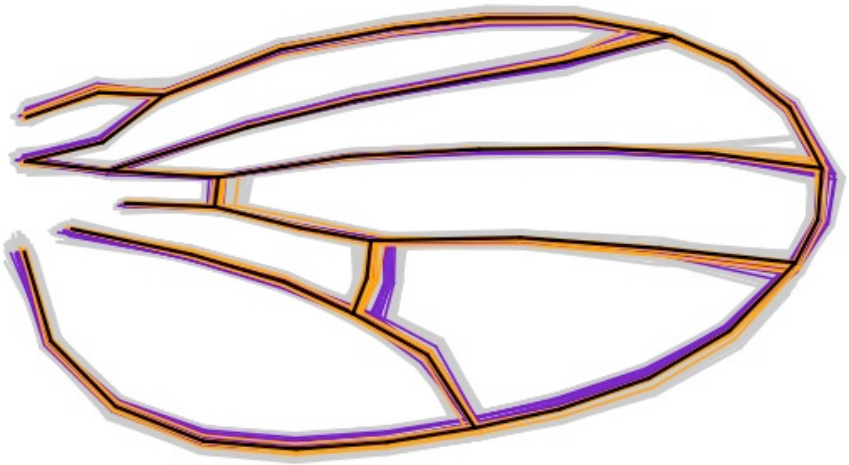
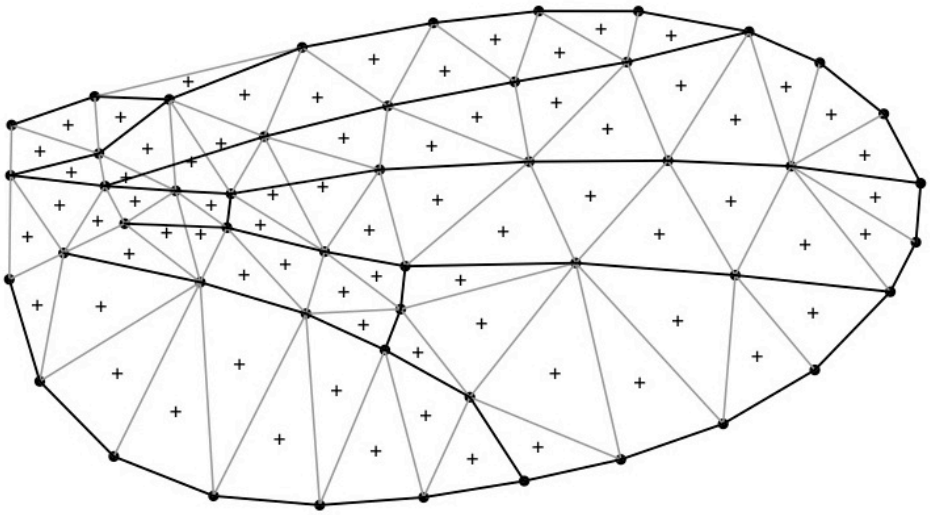


Figure 2: A, A representative wing image with fitted splines and the location of landmarks and semilandmarks along them. **B,** Individual variation in the sample (grey) relative to Sam (black), mutant means (orange), and means for naturally-derived lines (purple).

A



B

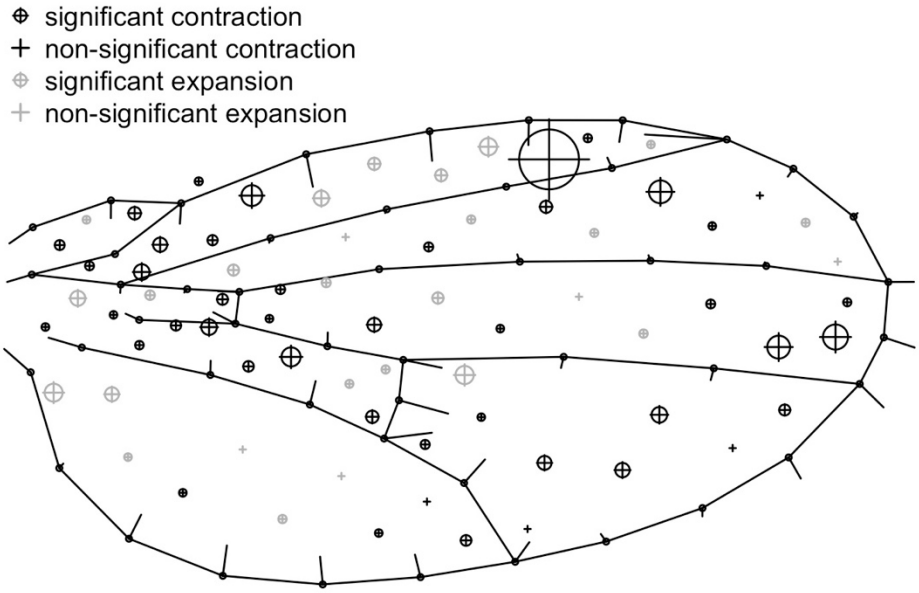


Figure 3: Illustrating the LORY method for interpolating shape changes from Procrustes data of genotype *Bs*^{k07909}. **A**, the tessellation scheme for choosing evaluation points. The landmarks and semi-landmarks (dots) are used as vertices. The centroids of the resulting triangles (crosses) are the final evaluation points. **B**, shape differences between the genotype mean and Sam inferred by calculating the Jacobians around each evaluation point, scaled by 100 (crosses and crossed circles). Bars represent the direction and change of differences between Sam and the sample mean at the landmarks and semi-landmarks based on the Procrustes data, scaled by 10X.

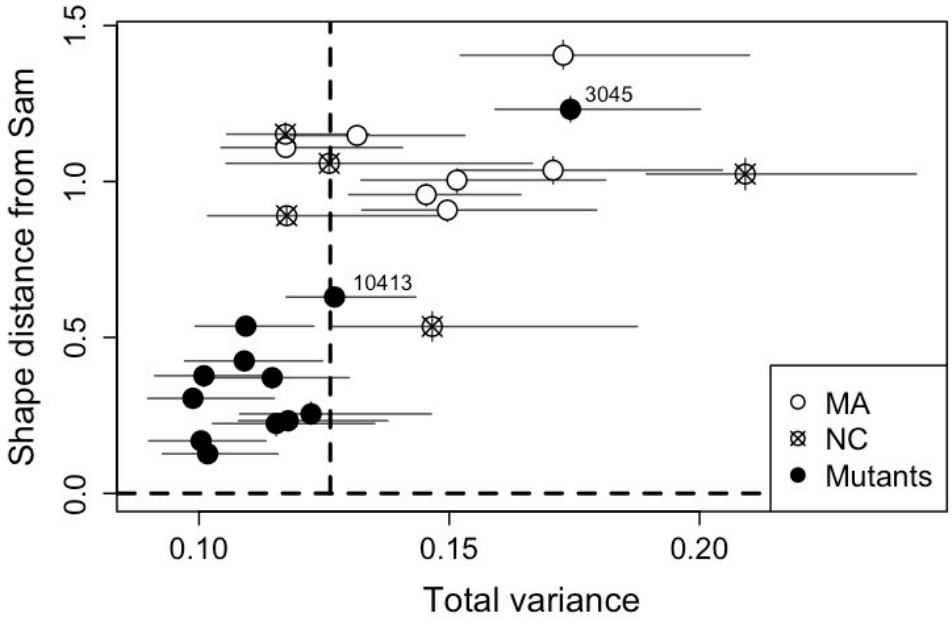


Figure 4: Mean-shape distance from Sam plotted against the total variance within each strain. Bars are 95% confidence intervals. Dashed lines represent values for Sam. Labeled mutants are mentioned in the text.

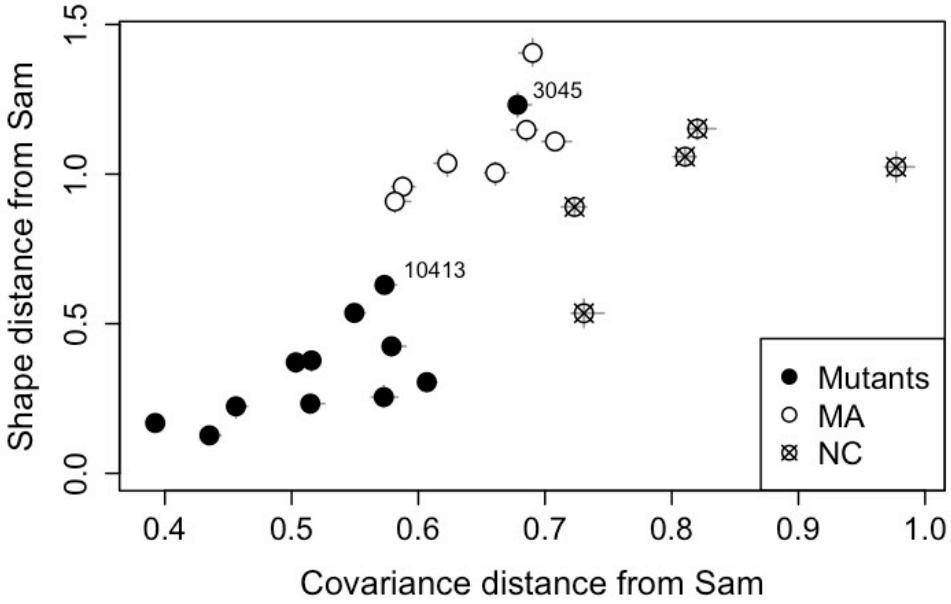


Figure 5: Mean-shape distance plotted against covariance distance (i.e, difference in matrix orientation; estimated by Random Skewers), relative to Sam. Bars are 95% confidence intervals. Labeled mutants are mentioned in the text.

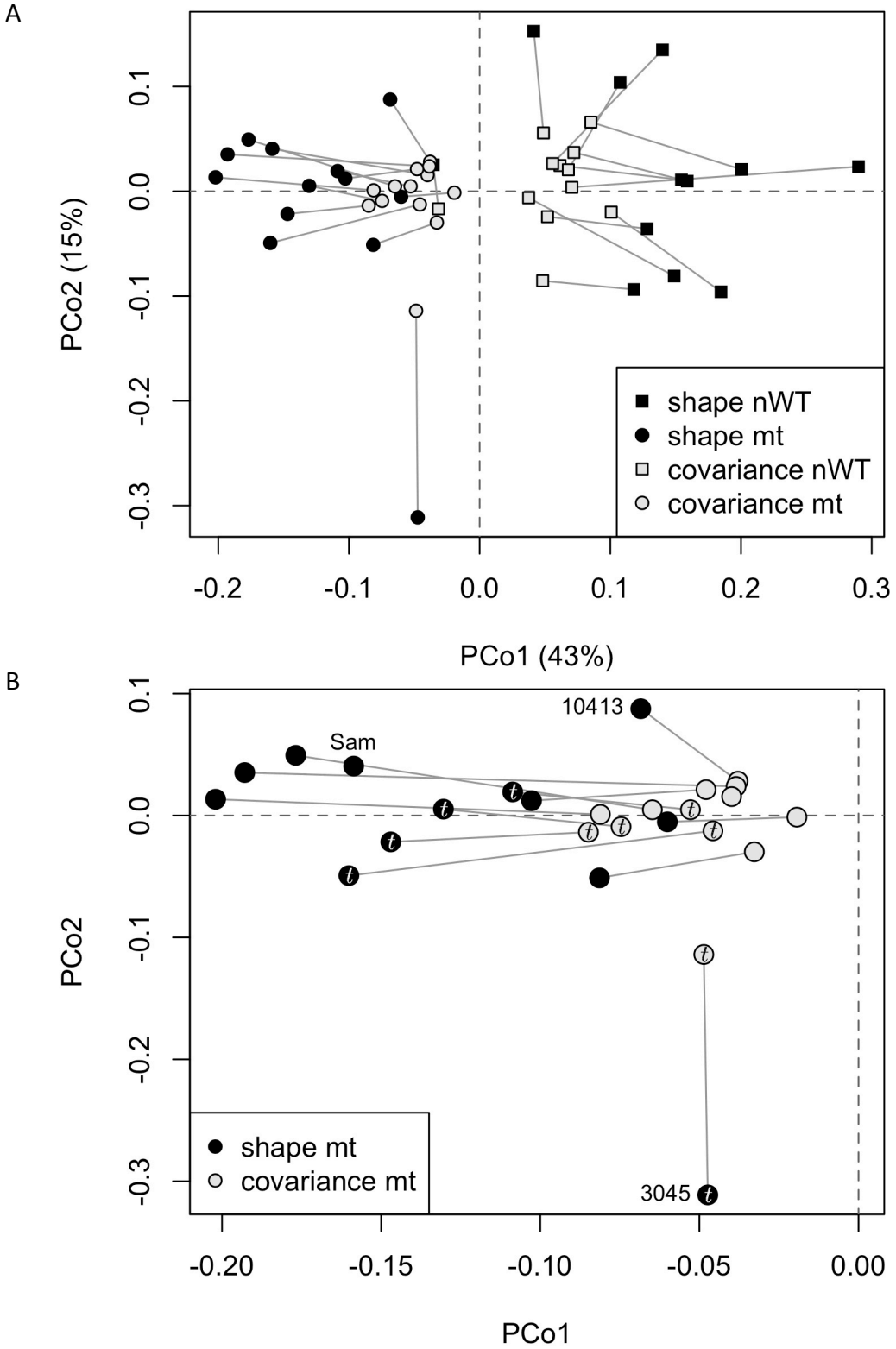


Figure 6: The first two axes of the covariance space superimposed onto shape space. Both spaces are based on all pairwise distances. Solid grey lines indicate the deviation between the two spaces for a given genotype. **A,** All genotypes; **B,** Mutants only, including Sam. Mutants that affect the TGF-β pathway are marked by *t*. All other mutants affect the EGFR pathway. nWT, naturally-derived wild types; mt, mutants. Labeled mutants are mentioned in the text.

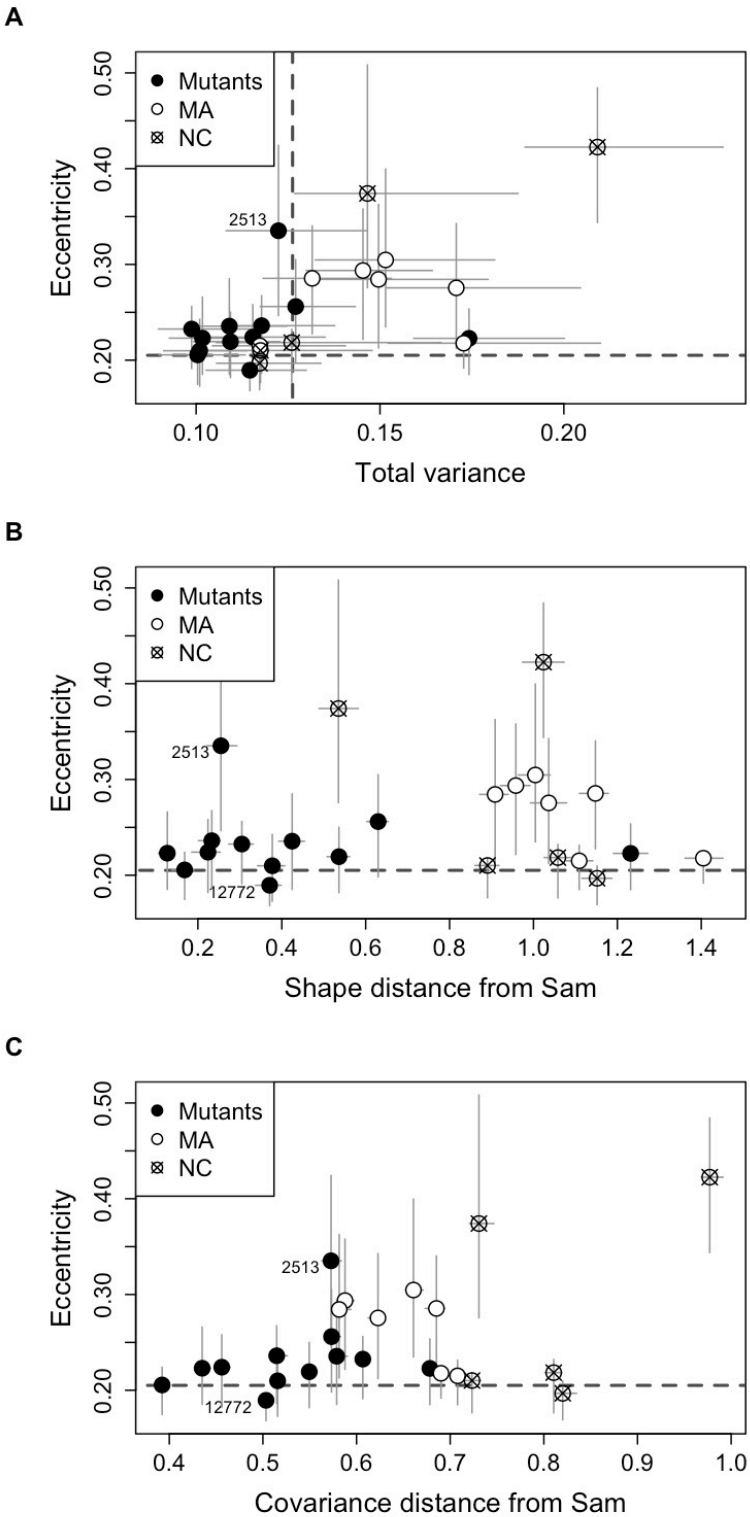


Figure 7: The association between eccentricity and (**A**) total variance within each genotype, (**B**) their mean-shape distance from Sam, and (**C**) their covariance distance from Sam. Dashed line represents the eccentricity of Sam. The apparent positive relationship between eccentricity and total variance is no greater than expected by chance based on parametric simulations. Labeled mutants are mentioned in the text.