The evolution of phenotypic integration: How directional selection reshapes covariation in mice

Article *in* Evolution · July 2017

DOI: 10.1111/evo.13304

5 authors, including:

Anna Penna
University of Texas at San Antonio
3 PUBLICATIONS **0** CITATIONS
SEE PROFILE

Diogo Melo
University of São Paulo
12 PUBLICATIONS **74** CITATIONS
SEE PROFILE

doi:10.1111/evo.13304

The evolution of phenotypic integration: How directional selection reshapes covariation in mice

Anna Penna,[1,2,*] Diogo Melo,[1,*] Sandra Bernardi,[3] Maria Inés Oyarzabal,[4] and Gabriel Marroig[1]

[1]*Laboratório de Evolução de Mamíferos, Departamento de Genética e Biologia Evolutiva, Instituto de Biociências, Universidade de São Paulo, Brazil*

[2]*E-mail: anna.penna@usp.br*

[3]*Cátedra de Histología y Embriología Básica. Facultad de Ciencias Veterinarias, Universidad Nacional de Rosario, Argentina*

[4]*Cátedra de Producción de Bovinos para Carne, Facultad de Ciencias Veterinarias y Consejo de Investigaciones, Universidad Nacional de Rosario, Argentina*

Received February 7, 2017
Accepted June 10, 2017

Variation is the basis for evolution, and understanding how variation can evolve is a central question in biology. In complex phenotypes, covariation plays an even more important role, as genetic associations between traits can bias and alter evolutionary change. Covariation can be shaped by complex interactions between loci, and this genetic architecture can also change during evolution. In this article, we analyzed mouse lines experimentally selected for changes in size to address the question of how multivariate covariation changes under directional selection, as well as to identify the consequences of these changes to evolution. Selected lines showed a clear restructuring of covariation in their cranium and, instead of depleting their size variation, these lines increased their magnitude of integration and the proportion of variation associated with the direction of selection. This result is compatible with recent theoretical works on the evolution of covariation that take the complexities of genetic architecture into account. This result also contradicts the traditional view of the effects of selection on available covariation and suggests a much more complex view of how populations respond to selection.

KEY WORDS: **Artificial selection, body size, cranium, experimental evolution, G-matrix, morphological integration, P-Matrix.**

Evolutionary change can only occur in the presence of variation, and when dealing with complex multivariate phenotypes (consisting of multiple traits) the patterns and magnitude of genetic covariation between traits can radically influence the course of evolution (Lande 1979; Felsenstein 1988). The standing genetic covariation of a given population depends on its evolutionary history, and can be altered by selection, drift, mutation, and recombination (Turelli and Barton 1994; Jones et al. 2004, 2014). These changes in covariation, in turn, can alter how a population responds to further selection or other evolutionary processes. So, if we are to understand how populations evolve and how the current phenotypic diversity observed in nature came to be, then the question of how genetic variation changes under various evolutionary processes becomes central to biology (Mitchell-Olds et al. 2007).

How a single trait responds to directional selection is a well-studied problem (Falconer and Mackay 1996). In general, we expect the response to selection to gradually erode genetic variation, as the many loci influencing a given trait go to fixation and, in the absence of mutation, preclude further evolutionary change (Bulmer 1971). If mutation is present and of sufficient magnitude, the variation removed by selection can be replenished and the rate of evolutionary change remains constant, at least for a time.

A theory on how directional selection and covariation interact to produce the response to selection on multiple traits was proposed by Lande (1979). This author related the standing genetic

*These authors contributed equally to this work.

covariation, represented by the additive genetic covariance matrix (G), to the selection gradient (β), a vector of selection coefficients acting independently on each trait, to predict the response to selection ($\Delta\bar{z}$). The Lande equation ($\Delta\bar{z} = G\beta$) predicts a response that is dependent on covariation, which means that selection can even lead to changes in traits that were not directly under selection (Cheverud 1984). Since more traits are presumably affected by more loci, in the multivariate case genetic architecture can become quite complicated, and consequently the erosion of genetic variation by selection is much more complex (Wolf et al. 2000; Pavlicev et al. 2008; Wagner and Zhang 2011). This conflict between selection and the maintenance of genetic variation in multiple traits remains a fundamental and puzzling problem in evolutionary biology (Walsh and Blows 2009).

One way to investigate the effects of selection on covariation is to use artificial selection experiments. Wilkinson et al. (1990) used experimental *Drosophila* populations selected for body size and found significant differences in correlation patterns between the different directions of selection, and even changes in the sign of some genetic correlations. Notwithstanding, analyzing the same dataset, Shaw et al. (1995) showed that the differences found by Wilkinson et al. (1990) are compatible with drift. Bryant et al. (1986) and Whitlock et al. (2002) also used *Drosophila* to show that both an increase and a decrease of genetic variation is possible under drift, probably due to genetic interactions like dominance and epistasis (Cheverud and Routman 1995). Taken together, these results present a conflicting picture on how covariation evolves, and on the potential evolutionary consequences of these changes.

On a macroevolutionary scale, retrospective studies have attempted to quantify the interplay of genetic constraints and phenotypic divergence. Pitchers et al. (2014) found no consistent pattern on how genetic covariation and intensities of selection affect the magnitude of evolutionary response. However, they did not take the multivariate aspect of the phenotype space into consideration and, since the orientation of selection alters the available variation for response, this can explain the lack of a clear pattern. As for the general pattern of bivariate correlations, Roff and Fairbairn (2012) surveyed estimates of selection gradients and correlations to test the hypothesis that high correlation between co-selected traits are advantageous, and found that traits that are selected in the same direction indeed tend to be more correlated than average. A well-studied case is the mammalian cranium, where covariation patterns tend to be stable (Porto et al. 2009), and phenotypic divergence is frequently size related and aligned with the main axis of within-population covariation, the genetic line of least resistance (Schluter 1996; Marroig and Cheverud 2010; Marroig et al. 2012; Porto et al. 2013). If this alignment of divergence and variation is a consequence of genetic constraints limiting the response to selection, or a case of selection altering covariation, or both, is still an open question.

Recently, several attempts have been made to directly investigate the behavior of the G-matrix under directional selection. Careau et al. (2015) used a large population of mice to select for changes in a multivariate behavior trait, and showed that selection reduced the available variation for adaptive response, significantly reducing the rate of adaptation after a few generations and reducing the amount of variation in the direction of selection. This is consistent with a traditional model of selection depleting variation extended to a multivariate context: even when variation is present in all traits in a complex system, combinations of traits may lack additive variation to respond to selection (Hine et al. 2011). On the other hand, Assis et al. (2016) used historical and modern samples of wild chipmunks separated by 100 generations, and showed that both the mean values of several cranium traits and the covariation between them had been altered by natural selection. Surprisingly, multivariate variation had increased in the direction of selection, suggesting that past selection can influence the standing covariation in a nonintuitive way, potentially facilitating further evolutionary responses in the same directions. These conflicting results regarding the interaction of multivariate variation and directional selection can be understood in light of recent theoretical work that allows for complex genetic architectures.

Heritable genetic covariation is determined by aspects of the genetic architecture like pleiotropy and linkage disequilibrium. Complex multivariate phenotypes, like skeletal traits, are influenced by a large number of loci, and covariation among these traits can be reasonably predicted from the pattern of shared pleiotropy (i.e., traits that have more loci affecting them simultaneously tend to be more correlated (Kenney-Hunt et al. 2008)). Porto et al. (2016) tested this hypothesis directly, comparing the level of pleiotropy between two species with different levels of phenotypic integration, and found that the more integrated species also showed higher levels of pleiotropy. In addition to pleiotropy, gene interactions (epistasis) can significantly influence covariation and alter patterns of pleiotropy (Wolf et al. 2005, 2006; Pavlicev et al. 2008). Simulations using the information that epistatic interactions can provide variation in pleiotropy and covariance patterns have shown how selection can promote changes in pleiotropy that can change covariation (Jones et al. 2014; Melo and Marroig 2015). Furthermore, Pavlicev et al. (2011) proposed a model for phenotypic evolution of multiple traits accounting for the influence of epistasis on the covariation between traits that shows how natural selection can increase the amount of variation along the direction of selection, exactly the kind of effect observed in Assis et al. (2016) and Roff and Fairbairn (2012).

In this article, we attempt to elucidate how directional selection interacts with and molds covariation using an experimental approach. Since selection on size is responsible for major morphological diversification (Baker et al. 2015), we used an evolutionary experimental approach in mice to investigate how evolutionary

changes in a multivariate system of phenotypic traits alters the pattern and magnitude of association between these traits. Mice lines were selected for an increase and for a decrease in overall size, and we focused on the evolutionary consequences of the changes in standing covariation in the cranium. Under a traditional additive model of covariation, we expect selection to deplete variation in the direction of selection, while under a more complex genetic architecture, including epistatic variation in pleiotropy, we expect the variation to be reorganized and increased in the direction of selection. This experimental approach allows us to understand the maintenance and reorganization of variation in a complex system and its consequences to evolution.

Materials and Methods

EXPERIMENTAL SELECTION LINES

This study was performed using animals from a long-term experiment involving one control line and two pairs of selected lines. In total, we used five groups: one control line t and four selected lines: upwards s', downwards s, upwards h', and downwards h. In 1985, a control population t with an effective population size $Ne \approx 40$ was founded with breeders chosen at random from an CF1 outbred population of $Ne \approx 80$ at the Facultad de Ciencias Veterinarias (Universidad Nacional de Rosario, Argentina). Then, two line-pairs of two-way individual selection for body weight at 49 days of age (s and h: downwards selected lines; s' and h': upwards selected lines) were founded from control line t, with mice drawn from generation three for s and s' and from generation eight for h and h' (Fig. S1, Oyarzabal 2011; Renny et al. 2014).

For control t line, effective population size was maintained by randomly choosing 20 individuals of each sex and avoiding full-sib mating. Selection on overall size was performed choosing the heaviest (for upwards lines) or lightest (for downwards lines) individuals for reproduction in each generation. Average effective population size of selected lines was maintained selecting six breeders of each sex for the downwards lines and four breeders of each sex for upwards lines. The difference in the number of breeders between upwards and downwards lines was due to the lower fertility of the lightest mice (Bernardi et al. 2009). Full-sib mating was also avoided in all the selected lines, except for the first generations of h and h' lines. For all matings, females were exposed to males in the ratio of 1:1. All the animals were chosen regardless of their inbreeding coefficients and the selection differentials. After ≈ 50 generations, the increase in inbreeding coefficients and the standard cumulative selection differentials were similar for selected lines (Table S1). See supporting information for additional information on number of weighted animals per generation (Fig. S2) and on average weight per generation (Fig. S3) for the full experiment.

SAMPLES

We had access to ≈ 65 individuals from around the 50th generation of each line (for a total sample of 329), with well-balanced sex ratios (see Table S2 for details). Mice were euthanized by cervical dislocation according to the American Veterinary Medical Association 2007 Guidelines on Euthanasia. We prepared all specimens in a dermestarium and removed their cranium and mandible. We collected 32 homologous anatomical landmarks in both sides of their cranium (Fig. S4, but see Cheverud (1995) and Garcia et al. (2014) for more details and the rationale for choosing these landmarks). To reduce measurement error, each anatomical marker was captured twice using a Microscribe MX 3D digitizer (Immersion Corporation, San Jose, California). We calculated a set of 35 euclidean distances (Fig. S4) between the landmarks using the average between both sides of the cranium for symmetrical distances, and the average between replicas of each individual. We opted for linear distances instead of following the current trend of using landmark data in a Geometric Procrustes Analysis (GPA) because GPA tends to disperse local variation and lead to misleading conclusions in regards to integration and modularity (see van der Linde and Houle (2009) and Márquez (2012) for details and possible solutions inside a landmark approach) and so that our results would be comparable to other assessments of integration in the mammalian cranium (Porto et al. 2009, 2013).

DIRECTION OF PHENOTYPIC DIVERGENCE

The multivariate phenotypic mean is a vector consisting of the mean of each cranial trait. To identify the direction in multivariate space for the phenotypic divergence between the two-way divergent selections, we calculated the vector of mean phenotypic divergence (δz). Each element of this vector was calculated as the difference between the pooled multivariate phenotypic mean of the lines selected for increase in weight and the pooled multivariate phenotypic mean of the lines selected for decrease in weight. We test if δz is indeed related to cranial size by comparing this vector with an isometric vector of equal loadings in all traits, which represents a direction of isometric size variation. High correlation between δz and the isometric vector indicates δz is related to cranial size. We also follow Mosimann (1970) and use the geometric mean of the cranial distances as a measure of overall isometric cranial size on an individual. This measure is highly correlated with the centroid size of cranial landmarks.

COVARIANCE MATRICES

Phenotypic and additive genetic covariance matrices (P- and G-matrices) for cranial traits of each line were obtained using a Bayesian sparse factor mixed model (BSFG). This is a robust method for estimating high dimensional G- and P-matrices from limited samples proposed by Runcie and Mukherjee (2013) and implemented in MATLAB (2013) by the authors. We removed

differences due to age, generation, and sex by using these groups as fixed effects in the mixed model. The models ran for 3000 iterations of burn in, followed by 100,000 iterations with thinning interval of 100. Convergence was assessed by inspecting trace and auto-correlation plots. From this model, we obtained a posterior distribution of 1000 covariance matrices for each line that summarizes the uncertainty in the estimation of respective covariance matrices, and we used this distribution in all posterior analysis to generate posterior distributions for all the calculated statistics. This allows us to take uncertainty into account when comparing the lines. Default priors had little effect on the covariance matrices, and mean posterior matrices were similar to matrices from a traditional MANCOVA.

Here we use P-matrices as a proxy for the respective additive genetic matrices (G-matrix), which are the important parameter in multivariate evolution. This was done because when calculating the effective sample size of our pedigree using the approximations from Raffa and Thompson (2016) we arrived at very low effective samples: around one sibpair for each individual line and around ten sibpairs for the full pedigree. This means that the pedigree for the sample we measured is such that G-matrices for each line are estimated with far too much uncertainty to be useful (even when using the BSFG model). Fortunately, P-matrices are probably a better estimate of the underlying genetic covariance pattern (Roff 1995; Marroig et al. 2012) and are informative on the underlying pleiotropic structure of genetic effects (Kenney-Hunt et al. 2008; Porto et al. 2016). To test the validity of this approximation, we calculated a pooled within-group P-matrix and a pooled within-group G-matrix estimated by the BSFG model using all of the individuals and controlling for differences in means between the lines. We then compare these matrices to confirm that the P- and G-matrices are similar. Matrix correlation between the pooled-within P- and pooled-within G-matrices was 0.95 for Random Skewers method (Cheverud and Marroig 2007, see Table S3). Matrix correlation between the posterior mean P-matrices of each line and the pooled-within G-matrix were all above 0.86 (Table S3), supporting the idea that G and P are similar. This hypothesis of similarity between P and G has been tested several times for this set of traits, and has been shown to be quite accurate in rodents (Garcia et al. 2014), and in mammals overall (Cheverud 1988; Porto et al. 2009, 2016, 2015; Marroig and Cheverud 2010; Hubbe et al. 2016).

MATRIX COMPARISONS

Since we are interested in changes in the influence of the covariance patterns on evolution, we assess the overall level of similarity between the covariance matrices for all the lines using two comparison methods that have immediate evolutionary interpretations, the Bayesian versions of the Random Skewers method and Krzanowski method proposed in Aguirre et al. (2014). The Random Skewers method (Cheverud and Marroig 2007) uses the Lande equation to simulate the response to random selection gradients, and the responses to the same selection gradient are then compared using vector correlations (the cosine of the angle between them) for the two matrices being compared. Aguirre et al. (2014) uses these random responses to identify directions in which the set of matrices being compared differ in their amount of available variation. So, this method allows us to explore differences in the distribution of variation in multiple directions of the phenotype space. The Krzanowski method (Krzanowski 1979) measures how similar the spaces spanned by the first several eigenvectors of the matrices being compared. For two matrices, the Krzanowski correlation is the mean of the squared vector correlations between all pairs of the first eigenvectors, usually the first $n/2 - 1$, where n is the dimensionality of the matrices. A Krzanowski correlation of 1 means the spanned spaces are exactly congruent, and a correlation of zero means the spaces are orthogonal. We can expand this method to several matrices by defining the **H** matrix (Krzanowski 1979):

$$H = \sum_{i=1}^{p} A_i A_i^t$$

where A_i is a column matrix containing the first $n/2 - 1$ eigenvectors of the i-th matrix being compared, p is the number of matrices being compared, and ^{t} denotes matrix transposition. The expectation of H is p times the covariance matrix of eigenvectors, and H/p tends to this covariance matrix for large p. The eigenvalues of **H** are bounded by p, and any eigenvalue equal to p indicates the associated eigenvector can be reconstructed by a linear combination of the eigenvalues included in the A_i matrices, and so is shared by all the matrices. An eigenvalue of less than p indicates that at least one of the A_i matrices can not span that eigenvector, and so the space is not completely shared between all the matrices. To test if our matrices share the same subspace, we follow Aguirre et al. (2014) and construct a randomized set of matrices under the assumption that all matrices are sampled from the same population (see the supporting information in Aguirre et al. (2014)). Observed and randomized eigenvalues of **H** are compared using posterior credibility intervals. If the randomized and observed eigenvalues of **H** are the same, we conclude the matrices share the same subspace.

EVOLUTIONARY STATISTICS

To assess the evolutionary consequences of the selection regimes on the covariance matrices, we calculated a series of evolutionarily informative statistics. In the following, **G** is an arbitrary covariance matrix, $\mathbf{G}^{-1}$ is the inverse of **G**, $tr(\mathbf{G})$ is the trace of **G**, $\lambda_i^{\mathbf{G}}$ is the i-th eigenvalue of **G**, $< \cdot, \cdot >$ represents the dot product between two vectors, $cos(\cdot, \cdot)$ is the cosine of the angle between two vectors (or their vector correlation), and $E[\cdot]_\beta$

represents the expected value over random β vectors with unit norm. (A) The magnitude of integration, calculated as the mean of the squared correlations between all traits. (B) The proportion of variation associated with the leading eigenvalue ($E1\% = \lambda_1^{\mathbf{G}}/tr(\mathbf{G})$). (C) The ability of the populations to respond in the direction of selection, calculated as the **mean flexibility** (*sensu* Marroig et al. 2009), which is given by the mean vector correlation between random selection gradients and their respective expected response to selection given the Lande equation ($\bar{f} = E[cos(\mathbf{G}\beta, \beta)]_\beta$). (D) The available variation for directional selection, calculated as the **mean evolvability** (Hansen and Houle 2008), which is given by the mean projection of the responses to the random selection gradients on these same selection gradient ($\bar{e} = E[< \mathbf{G}\beta, \beta >]_\beta$). The mean evolvability can also be calculated as the trace of the matrix being considered divided by the number of dimensions, so it is clearly a measure of total variation (Hansen and Houle 2008). Also, from the definition of the cosine between two vectors, flexibility in the direction of a given unit β can also be expressed as the ratio between evolvability in the direction of β and the norm of $G\beta$. We used a set of 1000 random selection gradients to calculate mean flexibility and mean evolvability. We then used these statistics to investigate how directional selection is affecting the evolutionary potential of each line.

We also evaluated how the direction of phenotypic divergence was related to the standing variation in the P-matrices. In this symmetrical directional selection case (i.e., where the two directions of selection are aligned but opposite) the direction of divergence δz describes the same direction as Δz. Furthermore, because our selection gradient is related to cranial size, the direction of phenotypic divergence is also a better predictor of the actual direction of selection than the selection gradient estimated indirectly from the response to selection and the G-matrix using the Lande equation (see supporting information for more details, and Marroig et al. (2012)). We used three directional metrics, (1) we calculated the ratio of the evolvability in the direction of the mean phenotypic divergence (with $\delta\hat{z}$ being the unit vector in the direction of divergence) and the mean evolvability along random phenotypic directions (**scaled directional evolvability**) as a measure of how biased the variation is in the direction of phenotypic divergence ($< \mathbf{G}\delta\hat{z}, \delta\hat{z} > /\bar{e}$); (2) similarly, we also measured the change in conditional evolvability, which measures the mean response to selection in the direction of a given β when other directions are under stabilizing selection (Hansen and Houle 2008). Mean conditional evolvability is calculated as $\bar{c} = E[(< \mathbf{G}^{-1}\beta, \beta >)^{-1}]_\beta$, and we define the ratio between conditional evolvability in the direction of phenotypic divergence and mean conditional evolvability along random phenotypic directions as the **scaled directional conditional evolvability** ($< \mathbf{G}^{-1}\delta\hat{z}, \delta\hat{z} >)^{-1}]/\bar{c}$) (3) we compared the alignment between δz and the first eigenvector of the covariance matrix for each line, using vector correlations. This correlation is a measure of how aligned the main axis of variation in the population is with regards to the realized evolutionary change. Since the first eigenvector in mammalian cranial matrices is usually related to size variation (Porto et al. 2009) and selection was on overall size, we expect a high alignment between the phenotypic divergence and standing variation. We test if the first eigenvectors (E1s) are indeed related to cranial size by comparing the E1s of the mean posterior matrices of all lines with the isometric size vector. High correlation between the E1s and the isometric vector indicates the E1s are related to cranial size.

Results

PHENOTYPIC DIVERGENCE

Both the upwards and the downwards selection resulted in changes in the weight of the corresponding s|s' and h|h' lines (Figs. S3 and S5). Both directional selection induced lower variation in weight, with the animals from the downwards selection showing smaller weights than those from the upwards selection, regardless of sex. In contrast, the control t line exhibited a large weight variation, which spanned the full variation range presented by the selected animals. The cranial traits also showed divergence between all four selected lines, with downwards selection showing smaller cranial size than those from the upwards selection, and the control t line being somewhat superimposed with the upwards lines (Figs. 1 and S6 for the complete set of cranial traits). The vector of phenotypic divergence (δz) had a correlation of 0.82 with the isometric vector, indicating that divergence in cranial traits was mainly in the cranial size direction.

MATRIX COMPARISONS

In the Krzanowski subspace comparisons all eigenvalues were not significantly different between the observed and randomized matrix comparisons (Fig. 2 A). As for the the Bayesian Random Skewers projection, we identified several directions with different amounts of phenotypic variation in each line (first few directions in Fig. 2 B and full results in Fig. S7). In most directions, the control line has significantly more variation than the selected lines. But this reduction is not uniform in all directions, and in several directions the selected lines have variation that is comparable to the control line. We provide in the supporting information the same set of results using G-matrices (Fig. S8). These results are considerably noisier, but consistent with the results obtained using the P-matrices.

EVOLUTIONARY STATISTICS

The distribution of mean values of evolutionary statistics showed a rise in integration and in the proportion of variation associated with the first eigenvector in selected lines when compared to the

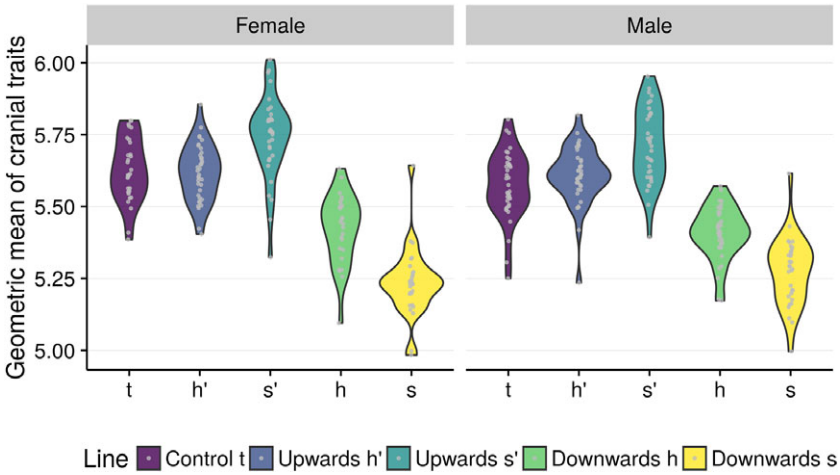

Figure 1. Distribution of the standardized isometric cranial size variable, calculated as the geometric mean of cranial traits by line and sex. Upwards lines are larger than downwards lines, while control line t is similar to the upwards lines.

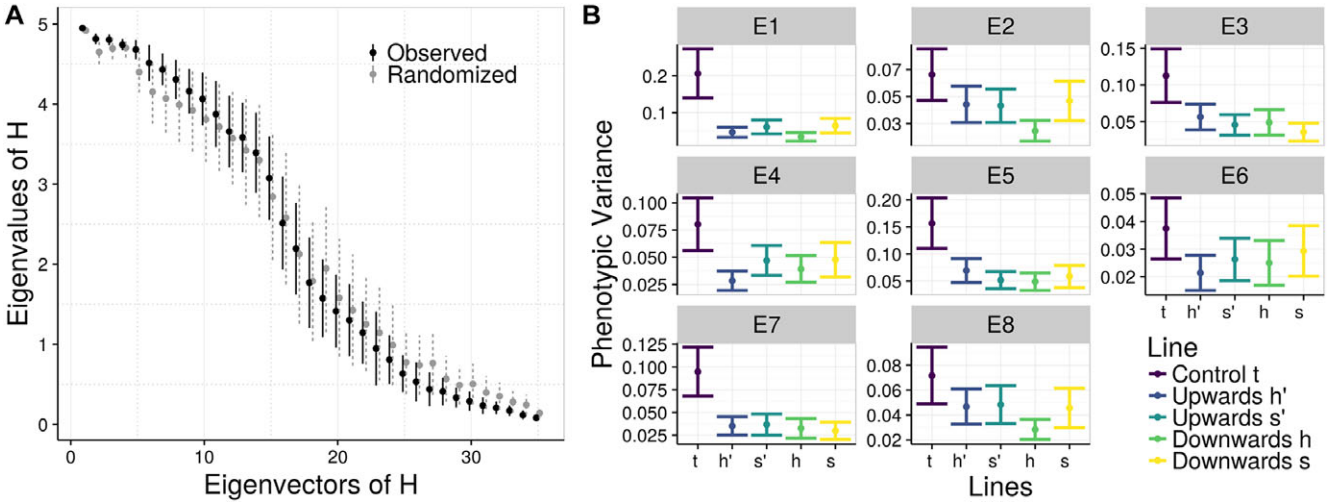

Figure 2. (A) Bayesian Krzanowski shared subspace. Eigenvalues for the H matrix are not significantly different in the randomized and observed matrices, indicating a shared subspace and a stable set of eigenvectors in all the lines. (B) Bayesian Random Skewers Projection. Here we only show the first eight eigenvectors of the decomposition, which are representative of the full set of eigenvectors (Fig. S7). In most directions the control line has higher variation than the selected lines, but in several directions the control and selected lines show comparable levels of variation, indicating that the loss of variation in the selected lines was not uniform in all directions. A version of this figure using the G-matrices is available in the Supporting Information (Fig. S8)

control t line (Fig. 3 A and B). All selected lines also showed a decline in evolvability and flexibility when compared to the control (Fig. 3 C and D). First eigenvectors for each line are given in Table S4, along with correlations between the first eigenvector and an isometric size vector. All correlations between E1s and the isometric vector were higher than 0.71, indicating that all first eigenvectors are related with cranial size.

The scaled directional evolvability, the ratio of evolvability in the direction of δz and the mean evolvability, shows a clear increase with selection, with selected lines showing about double the ratio observed in the control line (Fig. 4 A). The vector correlation between δz and the E1 increased in all selected lines, while this correlation in the control line shows a lower mean correlation (around 0.6) and wider posterior distribution. Downward selected lines have correlations of first eigenvector and δz above 0.9 and upwards selected lines above 0.75 (Fig. 4 B).

Discussion

Our experimental approach allows us to investigate how directional selection alters multivariate covariation, and what are the consequences of this interaction to evolution and diversification. Selected lines diverged in both weight and cranial traits, and the changes in the cranium were aligned with the main axis

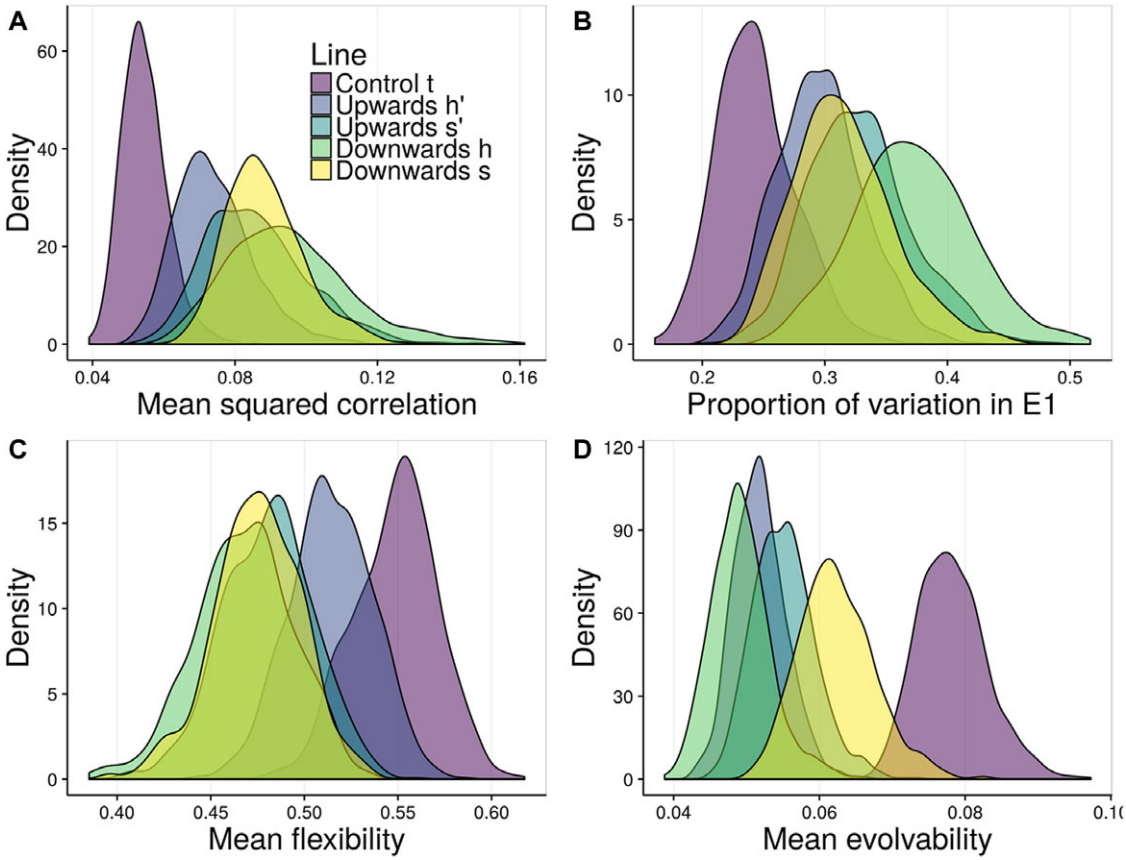

Figure 3. **Comparison of covariation patterns between control and selected lines using evolutionary statistics. (A) Magnitude of integration measured as mean squared correlation between all traits; (B) Proportion of variation associated with the first eigenvector, which is related to size variation; (C) Mean flexibility; (D) Mean evolvability. Curves represent posterior distributions obtained using the sample of P-matrices from the BSFG model. Confidence interval and mean for all curves can be found in Table S5. The same set of results using G-matrices are available in Fig. S9. G-matrix results are noisier, but consistent with the results obtained using the P-matrices, suggesting the changes we observed indeed occurred in the G-matrices.**

of variation, which is a direction associated with cranial size. While to some extent this is an unsurprising result given that selection was on overall size, our results illustrate how rapidly covariation can change under directional selection. In particular, the magnitude of association between traits is more malleable than the pattern of association.

Patterns of covariation were fairly similar in all lines, a result consistent with the observed pattern for natural populations of mammals in general, even when comparing covariation between different orders (Porto et al. 2009). Krzanowski subspace comparison showed all the lines share the subspace spanned by the first half of the eigenvectors. Regarding the distribution of variation, the Bayesian Random Skewers showed that in some directions the control line has more variation than the selected lines, while in others control and selected lines have comparable levels of variation. Taken together, these matrix comparison results indicate a stable set of eigenvectors, spanning a similar space in all the lines, and a nonisotropic reduction in variation in the selected lines, that is, some directions lost more variation than others (Figs. 2 and S7).

As for integration, all selected lines increased their magnitude of association between traits. The maximum observed difference in mean squared correlation is almost 0.05 (when comparing the control t line and the downwards h line) a difference comparable to those observed between mammalian orders (Marroig et al. 2009). This increase in integration is mirrored by the observed increase in the proportion of variation associated with the first eigenvector, which is related to size. This is expected, because size variation can be interpreted as coordinated variation in all traits, and so an increase in size variation leads to higher correlations between all traits (Porto et al. 2013). Therefore, the rise in integration in the selected lines is due to an increase in the proportion of variation that is related to cranial size (which changed due to selection on overall size).

This increase in integration is also reflected in the flexibility, which is lower in all the selected lines. Flexibility measures the

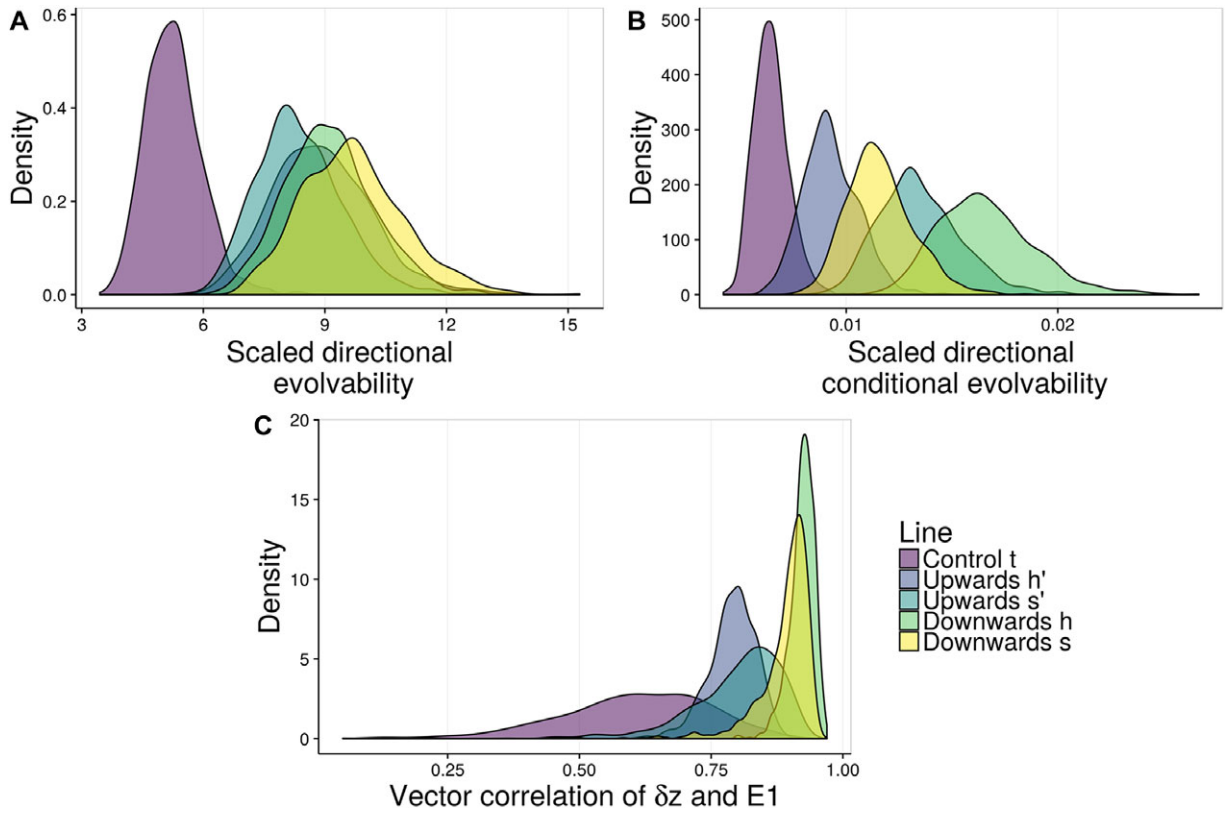

Figure 4. **(A) Scaled direction evolvability, the ratio of evolvability in the direction of phenotypic divergence (δz) and the mean evolvability for each line; (B) Scaled direction conditional evolvability, the ratio of conditional evolvability in the direction of δz and the mean conditional evolvability for each line; (C) the vector correlation between δz and first eigenvector for each line. Curves represent posterior distributions obtained using the sample of P-matrices from the BSFG model. Confidence interval and mean for all curves can be found in Table S6. The same set of results using G-matrices are available in Fig. S10. G-matrix results are noisier, but consistent with the results obtained using the P-matrices, suggesting the changes we observed indeed occurred in the G-matrices.**

ability of a population to respond in the direction of selection, and because a larger proportion of cranial variation in the selected lines is concentrated in size variation, less variation is available to respond in other directions. Also, the mean evolvability, or total available variation for responding to selection, is smaller in the selected lines.

At first sight, this is compatible with the idea that directional selection depletes variation. However, focusing only on the reduction of total variation is misleading. Taking the multivariate aspect of this system in consideration, the distribution of the available variation is also changing. Although total variation in all directions is smaller in the selected lines (Fig. 3, panel D), the increase in the proportion of variation associated with the first eigenvector (Fig. 3, panel B) (which is highly related to cranial size, Table S4) means the proportion of variation in the direction of evolutionary divergence is in fact increasing in the selected lines (Fig. 3, panels A and B), and the first eigenvector is also more aligned with the direction of divergence in the selected lines (Fig. 4 C). This reorganization of variation is supported by the change in the scaled directional evolvability and scaled conditional evolvability, which show that the scaled variation in the direction of divergence is almost doubled in selected lines when compared to the control (Fig. 4 A and B).

This relative increase in variation is incompatible with the traditional view of depletion of variation in the direction of selection due to the fixation of additive alleles. Under a purely additive model of cranial variation, as rare additive alleles with effects on size move to fixation, they could increase genetic variation for size in the process (Burger and Lynch 1995). In large populations this increase in variation under directional selection can be maintained by new mutations that arise and sweep to fixation. In principle we cannot rule out this possibility, but we are observing this effect after about 50 generations of strong directional selection (Table S1), after which we would expect the initial increase of variation to have disappeared as those initially rare alleles become fixed, and, because of our low effective population sizes, the influx of new mutations should be negligible. Also, increase in size variation due to additive alleles would be more likely to occur in traits that are controlled by a small number of loci (Burger and Lynch 1995; Jain and Stephan 2015), which is not likely to be the

case for the cranium (Leamy et al. 1999; Wolf et al. 2005; Porto et al. 2016). Furthermore, the genetic basis of morphological covariation is unlikely to be purely additive (Phillips et al. 2001; Whitlock et al. 2002). The changes are also not likely to be due to drift alone, as all the selected lines are consistent in their covariation patterns, and under drift we would expect a more random distribution of differences between the lines. However, epistatic interactions can provide a source of standing variation that allows the increase of variation in the direction of selection (Cheverud and Routman 1995; Wagner et al. 2007; Pavlicev et al. 2011; Melo and Marroig 2015), and can bias further mutations to be aligned with this direction (Jones et al. 2007, 2014).

Because genetic architectures can interact with selection in different ways, we expect different outcomes depending on the complexity of the genetic architecture underlying the set of traits under investigation: additive variation is expected to be consumed in the direction of selection, while epistatic variation can lead to an increase in variation in the direction of selection. Careau et al. (2015) showed that directional selection on behavior traits follows the expectation of the additive model, with a loss of variation in the direction of selection and a plateau in the response to selection after a few generations. In our experiment, as expected by the additive model, we also see a loss of total variation in the selected lines, but this reduction is nonisotropic. Some directions lost more variation than others, and the direction of selection ended up with proportionally more variation in the selected lines, suggesting the influence of epistasis and a more complex genetic architecture underlying the covariation of cranial traits. Assis et al. (2016) reported the same kind of realignment of variation and selection as we did, but saw no loss of variation in samples from modern populations after selection, when compared to historical samples. This could be related to the large effective sample sizes in the wild chipmunk populations considered, and suggests that large populations can reorganize covariation patterns in response to selection with no loss in total variation. Neither our results nor those from Assis et al. (2016) allow us to evaluate limits in the response to selection in the cranium, leading to plateaus in response, but we speculate that in large populations the availability of standing epistatic variation and the biasing of new mutations in the direction of selection could delay the onset of the kind of plateau in multivariate evolutionary response seen in Careau et al. (2015).

The observed increase in integration over a microevolutionary time scale also has consequences for our understanding of macroevolutionary patterns. In mammals, changes in the magnitude of integration are much more common than changes in the pattern of trait association (Porto et al. 2009). Our results suggest that these difference can be explained by the pervasiveness of selection on size along the mammalian clade (Marroig and Cheverud 2005; Baker et al. 2015). Lineages that underwent selection on size might have higher integration, and those whose selective response were not size-related might have lower integration. Also, divergence that is aligned with covariation can not be interpreted as only a product of constraints, since selection can directly reshape variation (Punzalan and Rowe 2016).

Here, we attempt to elucidate the effect of directional selection on covariation, and how this impacts evolution. We use experimental selection on size to answer this question, and find that selection can actively restructure covariation, and, in addition of depleting multivariate covariation, can reorganize standing covariation in the direction of evolutionary response, increasing a population's relative ability to respond to selection in that direction. This is in accordance with recent models of phenotypic covariation that include a more realistic genetic architecture, and an experimental evidence of directional selection shaping variation in nonintuitive ways. Along with recent empirical, theoretical, and simulation work, our results reinforce a shift in our understanding of how populations are shaped by natural selection. An obvious next step is to combine this sort of experimental selection with genetic mapping to understand, at the genomic level, how this reorganization of variation is taking place.

AUTHOR CONTRIBUTIONS

GM and MO designed the research. MO and SB conducted the selection experiments. AP cleaned and measured all animals. AP, DM and GM designed the analysis. AP and DM performed the analysis and wrote the initial draft. AP, DM, MO and GM reviewed and edited the final manuscript.

ACKNOWLEDGMENTS

Andrea Kaufmann-Zeh helped with the initial draft of this manuscript. Harley Sebastião and Barbara Costa helped with data acquisition. Julia Laterza Barbosa made the cranium illustrations showing the landmarks. Monique Simon, Guilherme Garcia, Fabio Machado, Jason Wolf, Michael Turelli and Günter Wagner provided helpful comments and discussion on the manuscript. Comments by Luis-Miguel Chevin, Mihaela Pavličev, William Pitchers, and one anonymous reviewer helped improve the final version of this manuscript. AP was funded by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) grant number 2013/06577-8, DM was funded by FAPESP grant number 2014/26262-4, GM was funded by FAPESP grant number 2011/14295-7.

DATA ARCHIVING

Code and data for performing all analysis are available from github (github.com/diogro/ratones, https://doi.org/10.5281/zenodo.815003), individual weight at 49 days and all cranial measurements, pedigree, posterior distribution, confidence intervals, and mean of all P- and G-matrices are archived in Dryad (https://doi.org/10.5061/dryad.5gr8r). Evolutionary statistics and matrix comparisons were calculated in R (R Core Team 2016) using the EvolQG package (Melo et al. 2015) (version 0.2–5).

LITERATURE CITED

Aguirre, J. D., E. Hine, K. McGuigan, and M. W. Blows. 2014. Comparing G: multivariate analysis of genetic variation in multiple populations. Heredity 112:21–29.

Assis, A. P. A., J. L. Patton, A. Hubbe, and G. Marroig. 2016. Directional selection effects on patterns of phenotypic (co)variation in wild populations. Proc. Biol. Sci. 283:20161615.

Baker, J., A. Meade, M. Pagel, and C. Venditti. 2015. Adaptive evolution toward larger size in mammals. Proc. Natl. Acad. Sci. USA 112:5093–5098.

Bernardi, S. F., G. Brogliatti, and M. I. Oyarzabal. 2009. Ovarian structure in mice lines selected for weight. Anat. Histol. Embryol. 38:200–203.

Bryant, E. H., S. A. McCommas, and L. M. Combs. 1986. The effect of an experimental bottleneck upon quantitative genetic variation in the housefly. Genetics 114:1191–1211.

Bulmer, M. G. 1971. The effect of selection on genetic variability. Am. Nat. 105:201–211.

Burger, R., and M. Lynch. 1995. Evolution and extinction in a changing environment: a quantitative-genetic analysis. Evolution 49:151–163.

Careau, V., M. E. Wolak, P. A. Carter, and T. Garland, Jr. 2015. Evolution of the additive genetic variance-covariance matrix under continuous directional selection on a complex behavioural phenotype. Proc. Biol. Sci. 282:20151119.

Cheverud, J. M. 1984. Quantitative genetics and developmental constraints on evolution by selection. J. Theor. Biol. 110:155–171.

Cheverud, J. M. 1988. A comparison of genetic and phenotypic correlations. Evolution 42:958–968.

Cheverud, J. M. 1995. Morphological integration in the saddle-back tamarin (*Saguinus fuscicollis*) cranium. Am. Nat. 145:63–89.

Cheverud, J. M., and G. Marroig. 2007. Research Article Comparing covariance matrices: random skewers method compared to the common principal components model. Genet. Mol. Biol. 30:461–469.

Cheverud, J. M., and E. J. Routman. 1995. Epistasis and its contribution to genetic variance components. Genetics 139:1455–1461.

Falconer, D. S., and T. F. C. Mackay. 1996. Introduction to quantitative genetics. 4th ed. Longmans Green, Harlow, Essex, U.K.

Felsenstein, J. 1988. Phylogenies and quantitative characters. Annu. Rev. Ecol. Syst. 19:445–471.

Garcia, G., E. Hingst-Zaher, R. Cerqueira, and G. Marroig. 2014. Quantitative genetics and modularity in cranial and mandibular morphology of *Calomys expulsus*. Evol. Biol. 41:619–636.

Hansen, T. F., and D. Houle. 2008. Measuring and comparing evolvability and constraint in multivariate characters. J. Evol. Biol. 21:1201–1219.

Hine, E., K. McGuigan, and M. W. Blows. 2011. Natural selection stops the evolution of male attractiveness. Proc. Natl. Acad. Sci. U S A 108:3659–3664.

Hubbe, A., D. Melo, and G. Marroig. 2016. A case study of extant and extinct Xenarthra cranium covariance structure: implications and applications to paleontology. Paleobiology 42:465–488.

Jain, K., and W. Stephan. 2015. Response of polygenic traits under stabilizing selection and mutation when loci have unequal effects. G3 5:1065–1074.

Jones, A. G., S. J. Arnold, and R. Bürger. 2004. Evolution and stability of the G-matrix on a landscape with a moving optimum. Evolution 58:1639–1654.

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

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

Kenney-Hunt, J. P., B. Wang, E. A. Norgard, G. Fawcett, D. Falk, L. S. Pletscher, J. P. Jarvis, C. Roseman, J. Wolf, and J. M. Cheverud. 2008. Pleiotropic patterns of quantitative trait loci for 70 murine skeletal traits. Genetics 178:2275–2288.

Krzanowski, W. J. J. 1979. Between-Groups Comparison of Principal Components. J. Am. Stat. Assoc. 74:703.

Lande, R. 1979. Quantitative genetic analysis of multivariate evolution, applied to brain: body size allometry. Evolution 33:402–416.

Leamy, L. J., E. J. Routman, and J. M. Cheverud. 1999. Quantitative trait loci for early- and late-developing skull characters in mice: a test of the genetic independence model of morphological integration. Am. Nat. 153:201–214.

van der Linde, K., and D. Houle. 2009. Inferring the nature of allometry from geometric data. Evol. Biol. 36:311–322.

Márquez, E. J., R. Cabeen, R. P. Woods, and D. Houle. 2012. The measurement of local variation in shape. Evol. Biol. 39:419–439.

Marroig, G., and J. Cheverud. 2010. Size as a line of least resistance II: direct selection on size or correlated response due to constraints? Evolution 64:1470–1488.

Marroig, G., and J. M. Cheverud. 2005. Size as a line of least evolutionary resistance: diet and adaptive morphological radiation in New World monkeys. Evolution 59:1128–1142.

Marroig, G., D. A. R. Melo, and G. Garcia. 2012. Modularity, noise, and natural selection. Evolution 66:1506–1524.

Marroig, G., L. T. Shirai, A. Porto, F. B. de Oliveira, and V. De Conto. 2009. The evolution of modularity in the mammalian skull II: evolutionary consequences. Evol. Biol. 36:136–148.

MATLAB. 2013. version 8.1.0.604 (R2013a). The MathWorks Inc., Natick, Massachusetts.

Melo, D., G. Garcia, A. Hubbe, A. P. Assis, and G. Marroig. 2015. Evolqg - an r package for evolutionary quantitative genetics [version 3; referees: 2 approved, 1 approved with reservations]. F1000Res. 4:925.

Melo, D., and G. Marroig. 2015. Directional selection can drive the evolution of modularity in complex traits. Proc. Natl. Acad. Sci. USA 112:470–475.

Mitchell-Olds, T., J. H. Willis, and D. B. Goldstein. 2007. Which evolutionary processes influence natural genetic variation for phenotypic traits? Nat. Rev. Genet. 8:845–856.

Mosimann, J. E. 1970. Size allometry: size and shape variables with characterizations of the lognormal and generalized gamma distributions. J. Am. Stat. Assoc. 65:930–945.

Oyarzabal, M. I. 2011. Líneas de ratones originales como modelos experimentales en genética y mejoramiento animal. J. Basic Appl. Genet. 22:0–0.

Pavlicev, M., J. M. Cheverud, and G. P. Wagner 2011. Evolution of adaptive phenotypic variation patterns by direct selection for evolvability. Proc. Biol. Sci. 278:1903–1912.

Pavlicev, M., J. P. Kenney-Hunt, E. A. Norgard, C. C. Roseman, J. B. Wolf, and J. M. Cheverud. 2008. Genetic variation in pleiotropy: differential epistasis as a source of variation in the allometric relationship between long bone lengths and body weight. Evolution 62:199–213.

Phillips, P. C., M. C. Whitlock, and K. Fowler. 2001. Inbreeding changes the shape of the genetic covariance matrix in *Drosophila melanogaster*. Genetics 158:1137–1145.

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. Philos. Trans. R Soc. Lond. B Biol. Sci. 369:20130252.

Porto, A., F. B. de Oliveira, L. T. Shirai, V. De Conto, and G. Marroig. 2009. The evolution of modularity in the mammalian skull I: morphological integration patterns and magnitudes. Evol. Biol. 36:118–135.

Porto, A., R. Schmelter, J. L. VandeBerg, G. Marroig, and J. M. Cheverud. 2016. Evolution of the genotype-to-phenotype map and the cost of pleiotropy in mammals. Genetics 204:1601–1612.

Porto, A., H. Sebastião, S. E. Pavan, J. L. VandeBerg, G. Marroig, and J. M. Cheverud. 2015. Rate of evolutionary change in cranial

morphology of the marsupial genus Monodelphis is constrained by the availability of additive genetic variation. J. Evol. Biol. 28:973–985.

Porto, A., L. T. Shirai, F. B. de Oliveira, and G. Marroig. 2013. Size variation, growth strategies, and the evolution of modularity in the mammalian skull. Evolution 67:3305–3322.

Punzalan, D., and L. Rowe. 2016. Concordance between stabilizing sexual selection, intraspecific variation, and interspecific divergence in Phymata. Ecol. Evol. 6:7997–8009.

R Core Team. 2016. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available at https://doi.org/www.R-project.org/.

Raffa, J. D., and E. A. Thompson. 2016. Power and effective study size in heritability studies. Stat. Biosci. 8:264–283.

Renny, M., N. B. Julio, S. F. Bernardi, C. N. Gardenal, and M. I. Oyarzabal. 2014. Estimation of genetic variability level in inbred CF1 mouse lines selected for body weight. J. Genet. 93:483–487.

Roff, D. A. 1995. The estimation of genetic correlations from phenotypic correlations: a test of Cheverud's conjecture. Heredity 74:481–490.

Roff, D. A., and D. J. Fairbairn. 2012. A test of the hypothesis that correlational selection generates genetic correlations. Evolution 66:2953–2960.

Runcie, D. E., and S. Mukherjee. 2013. Dissecting high-dimensional phenotypes with Bayesian sparse factor analysis of genetic covariance matrices. Genetics 194:753–767.

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

Shaw, F. H., R. G. Shaw, G. S. Wilkinson, and M. Turelli. 1995. Changes in genetic variances and covariances: G Whiz! Evolution 49:1260–1267.

Turelli, M., and N. H. Barton. 1994. Genetic and statistical analyses of strong selection on polygenic traits: what, me normal? Genetics 138:913–941.

Wagner, G. P., M. Pavlicev, and J. M. Cheverud. 2007. The road to modularity. Nat. Rev. Genet. 8:921–931.

Wagner, G. P., and J. Zhang. 2011. The pleiotropic structure of the genotype-phenotype map: the evolvability of complex organisms. Nat. Rev. Genet. 12:204–213.

Walsh, B., and M. W. Blows. 2009. Abundant genetic variation + strong selection = multivariate genetic constraints: a geometric view of adaptation. Annu. Rev. Ecol. Evol. Syst. 40:41–59.

Whitlock, M. C., P. C. Phillips, and K. Fowler. 2002. Persistence of changes in the genetic covariance matrix after a bottleneck. Evolution 56:1968–1975.

Wilkinson, G. S., K. Fowler, and L. Partridge. 1990. Resistance of genetic correlation structure to directional selection in *Drosophila melanogaster*. Evolution 44:1990–2003.

Wolf, J. B., E. D. Brodie III, and M. J. Wade. 2000. Epistasis and the Evolutionary Process. 1st ed. Oxford Univ. Press, USA.

Wolf, J. B., L. J. Leamy, E. J. Routman, and J. M. Cheverud. 2005. Epistatic pleiotropy and the genetic architecture of covariation within early and late-developing skull trait complexes in mice. Genetics 171:683–694.

Wolf, J. B., D. Pomp, E. J. Eisen, J. M. Cheverud, and L. J. Leamy. 2006. The contribution of epistatic pleiotropy to the genetic architecture of covariation among polygenic traits in mice. Evol. Dev. 8:468–476.

Associate Editor: L. Chevin
Handling Editor: P. Tiffin

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Figure S1: Scheme of relationship between mice lines used in the study.
Figure S2: Number of weighted animals per generation by line and sex.
Figure S3: Average weight at 49 days per generation by line and sex, for the full experiment.
Figure S4: Cranial anatomical landmarks used in this study and set of euclidean distances calculated between them, following Cheverud (1995).
Figure S5: Distribution of weight by line and sex.
Figure S6: Distribution of all inter-landmark distances used in this study.
Figure S7: Bayesian Random Skewers results for all 35 dimensions.
Figure S8: (A) Krzanowski shared subspace; (B) Bayesian Random Skewers.
Figure S9: Comparison of covariation patterns between control and selected lines using evolutionary statistics (A) Overall integration measured as mean square correlation between all traits; (B) Proportion of variation associated with the first eigenvector, which is related to size variation; (C) Mean exibility; (D) Mean evolvability.
Figure S10: (A) Scaled direction evolvability, the ratio of evolvability in the direction of δ_z and the mean evolvability for each line; (B) Scaled direction conditional evolvability, the ratio of conditional evolvability in the direction of δ_z and the mean conditional evolvability for each line; (C) the vector correlation between δ_z and first eigenvector for each line.
Table S1: Number of animals per generation, inbreeding coefficients and selection diffierentials by line.
Table S2: Sample sizes by line and sex.
Table S3: Comparison between the posterior mean G-matrix and all posterior median P-matrices and within groups P-matrix.
Table S4: First eigenvectors (E1) for the posterior median P-matrix for each line, and correlation of the E1s with an isometric size vector (a vector with equal loadings in all traits).
Table S5: 95% confidance interval and mean of the curves of posterior distributions obtained using the sample of P-matrices from the BSFG model from Figure 3.
Table S6: 95% confidance interval and mean of the curves of posterior distributions obtained using the sample of P-matrices from the BSFG model from Figure 4.
Data S1: