nature portfolio

https://doi.org/10.1038/s41559-021-01413-3

Supplementary information

Correlational selection in the age of genomics

In the format provided by the authors and unedited

Supplementary Material

1. Evolution of the G-matrix by correlational selection

In the bivariate case, the **G**-matrix can be represented as an ellipse containing 95% of the breeding values of the individuals in a population[1] (Fig. 3). If the two traits are strongly genetically correlated, the ellipse will be eccentric and oriented in a way that it is not parallel to either trait axis. Importantly, the long-axis of the **G**-matrix represents a genetic line of least resistance[2], the direction in phenotypic space which harbors the most genetic variance and along which the population most easily evolves.

In an attempt to describe the forces that affect standing variation in quantitative traits, Lande[2] imagined a population at equilibrium where the forces of selection and mutation are exactly canceled, such that there is no net evolution of the **G**-matrix:

$$\begin{aligned} \Delta\mathbf{G} = \mathbf{G\gamma G} - \mu\mathbf{M} = 0 \\ \mathbf{G\gamma G} = \mu\mathbf{M}, \end{aligned} \tag{S1}$$

where **M** is the mutation matrix, and linear selection is absent (because the population is on its optimum). This simple result suggests the shape of selection and mutation are expected to be aligned in well-adapted populations, and theoretical work has explored this in depth[3–6].

2. The role of g_{max} in modularity and plasticity

By focusing on the axis of maximum genetic variation, g_{max}, we can consider a range of questions that are applicable to any axis of variation.

- Does g_{max} capture strong functional interactions among traits? Or does it span modules?

o Does g_{max} coincide with the combination of traits that maximises fitness (i.e., align with selection)?

o Do we have evidence of what makes an axis of high genetic variance (allele frequencies or effect sizes), and whether pleiotropic alleles might have larger effects (e.g., are under stronger stabilizing selection).

o The evolutionary consequences of selection will depend on the relationship between the environmental and genetic covariances of traits with fitness. Environment and genes might act through the same developmental processes, resulting in strongly concordant phenotypic and genetic correlations[7–10]. Thus, the distribution of phenotypic variation available for selection to act on reflects the distribution of variation that can evolutionarily respond to the selection.

o Recent theories of evolutionary responses to rapid environmental change (evolutionary rescue) predict that changes in variance among environments due to variation in plastic response to environments could accelerate adaptation to new environments[11,12]. Phenotypic plasticity, environmentally induced phenotypic variation, has long been predicted to facilitate population survival under both in situ environment change and colonisation, allowing individuals to maintain relatively high fitness by changing their phenotype toward the optimal value in the new environment[13]. In common with adaptation, plastic responses to environmental variation might typically involve changes in multiple traits, including different life history and pigmentation traits[14]. The role of plasticity and evolution in population survival are difficult to disentangle, as both plasticity and trait means can evolve rapidly, and few studies are able to map the phenotypic optimum in multivariate trait space, and measure both the initial plasticity and any rapid evolution of the trait mean and plasticity. Several studies have noted that the most

variable axes of multivariate phenotypic plasticity align with axes of phenotypic evolutionary divergence among taxa, suggesting that plasticity might have initially contributed to that evolution[15]. A meta-analysis of several studies reported that the major axis of quantitative genetic variance was typically well aligned with the plastic change in population multivariate phenotypic mean in response to a novel environment[16], suggesting that selection on plasticity and trait means would have synergistic effects on fitness. This question, of how multivariate selection acts to increase fitness under novel, or markedly altered environmental conditions, can be tackled with genomic data. For traits that do respond plastically, our understanding of the evolutionary consequences of plasticity and multivariate selection within an environment could be advanced by determining whether loci contributing to variation in the multivariate plastic response either have pleiotropic effects on the multivariate trait within an environment, or are in linkage disequilibrium with loci that do.

3. Supplemental material for Fig. 3

Figure 3A: Functional form of fitness surfaces

The fitness surfaces in Fig. 3A were produced using the following functions:

Additive selection

First column:

$$w(z_1, z_2) = K_1(e^{z_1} + e^{z_2})$$

Second column: (S2)

$$w(z_1, z_2) = K_2\left(e^{-z_1^2} + e^{-\left(z_2 - \frac{3}{2}\right)^2}\right)$$

Correlational selection

Third column:

$$w(z_1, z_2) = K_3 \left(\frac{e^{-\frac{1}{10}(z_1 - z_2)^2}}{1 + e^{-\frac{1}{2}(z_1 + z_2)}} \right)$$

Fourth: column:

$$w(z_1, z_2) = K_4 \left(e^{-\frac{1}{2}[(z_1+1)^2 + (z_2-1)^2]} + e^{-\frac{1}{2}[(z_1-1)^2 + (z_2+1)^2]} \right)$$

The constants K_i were chosen to ensure that $\mathbb{E}w(z_1, z_2) = 1$ when z_1 and z_2 are independent standard normal traits, which occurs when $K_1 = \frac{1}{2\sqrt{e}}$, $K_2 = \frac{\sqrt{3}}{1+e^{-3/4}}$, $K_3 \approx 2.37$ and $K_4 = \sqrt{e}$.

Figure 3B: Estimation of fitness surfaces

Here we provide further details about the fitness surfaces in Fig. 3B and their estimation using both quadratic regression and non-parametric methods. Note that several alternative definitions of the quadratic selection gradients γ co-exist in the literature, all of which originated in the work of Lande and Arnold[17]. These definitions coincide if traits are multivariate normal. For the non-parametric methods, we rely on a definition in terms of partial derivatives of the fitness surface (see ref. [18] for details).

True trait distributions and fitness surfaces

For all three rows of Fig. 3B, the traits z_1 and z_2 were assumed to be independent, with each following a standard normal distribution. The true fitness surfaces in the first column of Fig. 3B are given explicitly by the following functions:

Unimodal surface:

$$w(z_1, z_2) = K_5 \, e^{-\frac{1}{2}(z_1^2 + z_2^2)} \tag{S3}$$

Bimodal surface:

$$w(z_1, z_2) = K_6 \left(e^{-\frac{1}{2}[(z_1+1)^2+(z_2-1)^2]} + e^{-\frac{1}{2}[(z_1-1)^2+(z_2+1)^2]} \right)$$

Rising fitness ridge:

$$w(z_1, z_2) = K_7 \left(\frac{e^{-\frac{1}{10}(z_1-z_2)^2}}{1 + e^{-\frac{1}{2}(z_1+z_2)}} \right)$$

The constants K_i were chosen to ensure that $\mathbb{E}w(z_1, z_2) = 1$ when z_1 and z_2 are independent standard normal traits, which occurs when $K_5 = 2$, $K_6 = \sqrt{e}$ and $K_7 \approx 2.37$. The true values of γ_{12}^* for these fitness surfaces were calculated by numerical integration:

$$\gamma_{12}^* = \iint \frac{\partial^2 w(z_1, z_2)}{\partial z_1 \partial z_2} f(z_1, z_2)\, dz_1 dz_2, \qquad \text{(S4)}$$

where $f(z_1, z_2) = \frac{1}{2\pi} e^{-\frac{1}{2}(z_1^2+z_2^2)}$ is the joint probability density function for independent standard normal traits z_1 and z_2 (ref. [18]). Very similar estimates of γ_{12}^* are obtained by averaging $\frac{\partial^2 w(z_1,z_2)}{\partial z_1 \partial z_2}$ over the sample distributions of z_1 and z_2.

Simulation of sample data

The sample datasets in the second column of Fig. 3B were simulated as follows. First, trait values of z_1 and z_2 were drawn from independent standard normal distributions for $n = 1000$ individuals. Second, the fitness of each individual was simulated as a Poisson random variable with expected value $w(z_1, z_2)$ taken from equation (S3). Trait values were then centered to have means of zero, and fitness was normalized to have a mean of one.

Estimation using quadratic regression

Quadratic estimates of the fitness surface using the Lande-Arnold approach[17] are illustrated in the third column of Fig. 3B . They were obtained using linear models in R with the syntax w~z1+z2+I(z1^2)+I(z2^2)+I(z1*z2). The linear selection gradients β_1 and β_2 correspond to the regression coefficients for z1 and z2 respectively; the quadratic selection gradient γ_{12} corresponds to the coefficient for I(z1*z2); and the quadratic selection gradients γ_{11} and γ_{22} correspond to *twice* the regression coefficients for I(z1^2) and I(z2^2) respectively[19]. The full results of this analysis are presented in Table S1.

Although the values of z_1 and z_2 in Fig. 3B were sampled from a bivariate normal distribution, the sample distribution of the traits is not quite normal. Consequently, the estimates of $\boldsymbol{\gamma}$ obtained via quadratic regression (Table S1) differ slightly from those obtained using the primary definition $\boldsymbol{\gamma} = \boldsymbol{P}^{-1}\boldsymbol{C}\boldsymbol{P}^{-1}$ of Lande and Arnold (equation 14a in ref. [17]). These two definitions coincide only for multivariate normal traits.

Table S1: Full results of the quadratic regression underlying the third column of Fig. 3B. Results in bold are significant at $\alpha = 0.05$.

	Linear selection gradients		**Quadratic selection gradients**		
Fitness surface	β_1	β_2	γ_{11}	γ_{22}	γ_{12}
Unimodal	−0.02	0.03	**−0.51**	**−0.50**	−0.01
Bimodal	−0.02	−0.01	**−0.21**	**−0.24**	**−0.23**
Fitness ridge	**0.18**	**0.26**	**−0.16**	**−0.16**	**0.12**

Non-parametric estimates

Non-parametric estimates of the fitness surface are illustrated in the fourth column of Fig. 3B. They were obtained using thin-plate splines in the R package mgcv[20] using the syntax w~s(z1)+s(z2)+ti(z1,z2). The ti command explicitly estimates the interaction effect, corresponding to $I(z_1, z_2)$ in equation (B4) of Box 1. We write $\widehat{w}(z_1, z_2)$ for the non-parametric estimate of the fitness surface.

After obtaining estimates of the fitness surface, we estimated the average mixed derivatives $\gamma_{12}^* = \mathbb{E}\frac{\partial^2 w(z_1,z_2)}{\partial z_1 \partial z_2}$. Assuming that the second partial derivatives of the true fitness surface are continuous, the true value of the mixed derivative $\frac{\partial^2 w(z_1,z_2)}{\partial z_1 \partial z_2}$ at any point on the fitness surface $w(z_1, z_2)$ is given by:

$$\frac{\partial^2 w(z_1, z_2)}{\partial z_1 \partial z_2} = \lim_{h_1, h_2 \to 0} \frac{w(z_1 + h_1, z_2 + h_2) - w(z_1 + h_1, z_2) - w(z_1, z_2 + h_2) + w(z_1, z_2)}{h_1 h_2}. \quad \text{(S5)}$$

In practice, we estimated this as:

$$\frac{\partial^2 w(z_1, z_2)}{\partial z_1 \partial z_2} \approx \frac{\widehat{w}(z_1 + \Delta, z_2 + \Delta) - \widehat{w}(z_1 + \Delta, z_2) - \widehat{w}(z_1, z_2 + \Delta) + \widehat{w}(z_1, z_2)}{\Delta^2}, \quad \text{(S6)}$$

where Δ is a small constant (we used $\Delta = 0.001$). To estimate γ_{12}^*, we then averaged the point estimates in equation (S6) over the empirical distribution of z_1 and z_2. Note that in Fig. 3B, the traits were drawn from a multivariate normal distribution, and so the estimates of γ_{12}^* using equation (S6) are similar to estimates of γ_{12} using the Lande-Arnold approach. However, if traits are not multivariate normal, then these two estimates may not coincide.

Supplemental references

1. Steppan, S. J., Phillips, P. C. & Houle, D. Comparative quantitative genetics: evolution of the G matrix. *Trends Ecol. Evol.* **17**, 320–327 (2002).
2. Lande, R. The genetic covariance between characters maintained by pleiotropic mutations. *Genetics* **94**, 203–215 (1980).
3. Jones, A. G., Arnold, S. J. & Burger, R. Stability of the G-matrix in a population experiencing pleiotropic mutation, stabilizing selection, and genetic drift. *Evolution* **57**, 1747–1760 (2003).
4. Jones, A. G., Arnold, S. J. & Burger, R. Evolution and stability of the G-matrix on a landscape with a moving optimum. *Evolution* **58**, 1639–1654 (2004).
5. Jones, A. G., Arnold, S. J. & Burger, R. The mutation matrix and the evolution of evolvability. *Evolution* **61**, 727–745 (2007).
6. Jones, A. G., Burger, R. & Arnold, S. J. Epistasis and natural selection shape the mutational architecture of complex traits. *Nat.Commun.* **5**, 3709 (2014).
7. Cheverud, J. M. A comparison of genetic and phenotypic correlations. *Evolution* **42**, 958–968 (1988).
8. Roff, D. A. The estimation of genetic correlations from phenotypic correlations: a test of Cheveruds conjecture. *Heredity* **74**, 481–490 (1995).
9. Kruuk, L. E. B., Slate, J. & Wilson, A. J. New answers for old questions: the evolutionary quantitative genetics of wild animal populations. *Annu. Rev. Ecol. Evol. Syst.* **39**, 525–548 (2008).
10. Sodini, S. M., Kemper, K. E., Wray, N. R. & Trzaskowski, M. Comparison of genotypic and phenotypic correlations: Cheverud's conjecture in humans. *Genetics* **209**, 941–948 (2018).

11. Lande, R. Adaptation to an extraordinary environment by evolution of phenotypic plasticity and genetic assimilation. *J. Evol. Biol.* **22**, 1435–1446 (2009).

12. Ashander, J., Chevin, L.-M. & Baskett, M. L. Predicting evolutionary rescue via evolving plasticity in stochastic environments. *Proc. R. Soc. B Biol. Sci.* **283**, 20161690 (2016).

13. Price, T. D., Qvarnstrom, A. & Irwin, D. E. The role of phenotypic plasticity in driving genetic evolution. *Proc. R. Soc. Lond. Ser. B-Biol. Sci.* **270**, 1433–1440 (2003).

14. van Bergen, E. *et al.* Conserved patterns of integrated developmental plasticity in a group of polyphenic tropical butterflies. *BMC Evol. Biol.* **17**, 59 (2017).

15. Wund, M. A., Baker, J. A., Clancy, B., Golub, J. L. & Foster, S. A. A test of the "flexible stem" model of evolution: ancestral plasticity, genetic accommodation, and morphological divergence in the threespine stickleback radiation. *Am. Nat.* **172**, 449–462 (2008).

16. Noble, D. W. A., Radersma, R. & Uller, T. Plastic responses to novel environments are biased towards phenotype dimensions with high additive genetic variation. *Proc. Natl. Acad. Sci.* **116**, 13452–13461 (2019).

17. Lande, R. & Arnold, S. J. The measurement of selection on correlated characters. *Evolution* **37**, 1210–1226 (1983).

18. Walsh, B. & Lynch, M. *Evolution and Selection of Quantitative Traits*. (Oxford University Press, 2018).

19. Stinchcombe, J. R., Agrawal, A. F., Hohenlohe, P. A., Arnold, S. J. & Blows, M. W. Estimating nonlinear selection gradients using quadratic regression coefficients: double or nothing? *Evolution* **62**, 2435–2440 (2008).

20. Wood, S. N. *Generalized Additive Models: An Introduction with R*. (CRC Press, 2017).