Evolution of the G Matrix under Nonlinear Genotype-Phenotype Maps

Milocco, Lisandro

2022-03

Milocco , L & Salazar-Ciudad , I 2022 , ' Evolution of the G Matrix under Nonlinear Genotype-Phenotype Maps ' , American Naturalist , vol. 199 , no. 3 , pp. 420-435 . https://doi.org/10.1086/717814

http://hdl.handle.net/10138/355435
https://doi.org/10.1086/717814

Evolution of the *G* Matrix under Nonlinear Genotype-Phenotype Maps

Lisandro Milocco[1,*] and Isaac Salazar-Ciudad[1,2,*]

1. Institute of Biotechnology, University of Helsinki, Helsinki, Finland; 2. Centre de Recerca Matemàtica, Barcelona, Spain; and Genomics, Bioinformatics, and Evolution, Departament de Genètica i Microbiologia, Universitat Autònoma de Barcelona, Barcelona, Spain

Submitted February 11, 2021; Accepted September 3, 2021; Electronically published January 11, 2022

Online enhancements: supplemental PDF. *Dryad data:* https://doi.org/10.5061/dryad.z34tmpgck.

ABSTRACT: The *G* matrix is a statistical summary of the genetic basis of a set of traits and a central pillar of quantitative genetics. A persistent controversy is whether *G* changes slowly or quickly over time. The evolution of *G* is important because it affects the ability to predict, or reconstruct, evolution by selection. Empirical studies have found mixed results on how fast *G* evolves. Theoretical work has largely been developed under the assumption that the relationship between genetic variation and phenotypic variation—the genotype-phenotype map (GPM)—is linear. Under this assumption, *G* is expected to remain constant over long periods of time. However, according to developmental biology, the GPM is typically complex and nonlinear. Here, we use a GPM model based on the development of a multicellular organ to study how *G* evolves. We find that *G* can change relatively fast and in qualitative different ways, which we describe in detail. Changes can be particularly large when the population crosses between regions of the GPM that have different properties. This can result in the additive genetic variance in the direction of selection fluctuating over time and even increasing despite the eroding effect of selection.

Keywords: quantitative genetics, *G* matrix, genotype-phenotype map, evo-devo, mathematical modeling.

Introduction

Variation is a requisite for evolution. It is on the existing phenotypic variation that natural selection acts. Understanding the properties of phenotypic variation is therefore key to explaining evolutionary dynamics. In the framework of quantitative genetics, the properties of variation of a set of polygenic traits in a population are summarized by variance and covariance matrices. The most important of these are *P* and *G*, the phenotypic and additive-genetic (co)variance matrices, respectively (Falconer and Mackay 1996; McGuigan 2006).

The *G* matrix is a statistical summary of the additive genetic relationships between traits. The diagonal elements of *G* are the additive genetic variances of the traits, and the off-diagonal elements are the additive genetic covariances between traits. Additive genetic covariation arises, for example, from traits sharing part of their genetic, developmental, and physical bases. Because of these shared bases, the evolution of these traits is entangled: direct selection on one trait can lead to indirect selection on covarying traits. This is most clearly summarized by the multivariate breeder's equation (Lande 1979), $\Delta\bar{z} = G\beta$, which provides a prediction of the response to selection of trait means ($\Delta\bar{z}$) as the product of *G* and β, the vector of partial regression coefficients of relative fitness on traits.

A persistent controversy in quantitative genetics is whether the *G* matrix changes slowly or quickly over time (Steppan et al. 2002; Arnold et al. 2008; Aguirre et al. 2014). The *G* matrix can be used to reconstruct past natural selection (e.g., Merilä et al. 1994) and to predict the response to selection and, thus, future evolution. Estimating *G* is costly and time-consuming. If *G* changes quickly, the *G* estimated at a given time can lead to inaccurate predictions of the response to selection at other times. In other words, change reduces the predictive value of *G* (Eroukhmanoff 2009). Moreover, knowledge of the ways in which *G* can change is important to understanding how future evolution will proceed.

Empirical studies of the evolution of *G* have found mixed results for whether *G* changes quickly or slowly. Some comparisons of *G* matrices between natural populations found no evidence of change in *G* even for hundreds of generations (Delahaie et. al. 2017; Hangartner et al. 2020; others reviewed in Arnold et al. 2008), while others

* Corresponding authors; email: lisandro.milocco@helsinki.fi, isaac.salazar@uab.cat.

ORCIDs: Milocco, https://orcid.org/0000-0003-3953-0407; Salazar-Ciudad, https://orcid.org/0000-0003-1607-0962.

have found significant change in *G*, even in only a few generations (Cano et al. 2004; Doroszuk et al. 2008; Eroukhmanoff and Svensson 2011; Wood and Brodie 2015; Walter et al. 2018; Chakrabarty and Schielzeth 2020). Artificial selection experiments in controlled settings also show contrasting results (Shaw et al. 1995; Blows and Higgie 2003; Phelan et al. 2003; Careau et al. 2015; Penna et al. 2017). An important limitation of these studies is that because of the difficulties of estimating *G*, there are usually only a few estimates of *G* available, generally only two. Moreover, these estimates are typically associated with large sampling variance, making comparisons difficult and contributing to the inconsistency of empirical results.

The *G* matrix is a statistical abstraction, estimated from the phenotype without direct knowledge of the underlying gene networks and developmental biology. Theoretical research on the evolution of *G* follows a similar approach: the gene product interactions, gene networks, and epigenetic factors underlying phenotypes and their variation are not explicitly considered. Indeed, models are purposely kept simple for analytical tractability. Thus, the relationship between the genetic variation and the phenotypic variation, or the genotype-phenotype map (GPM), has typically been assumed to be linear (Lande 1979; Turelli 1985; Jones et al. 2003, 2012) or to include certain regular patterns of epistasis (Jones et al. 2014).

The complexity and structure of the GPM is of particular interest for the field of evolutionary developmental biology. Research from this field suggests that the GPM is typically highly nonlinear and complex and that this has important consequences for evolutionary dynamics in general (Alberch 1982; Raff 1996; Newman and Müller 2000; Müller 2007) and for evolutionary quantitative genetics in particular (Houle 1991; Rice 2004, 2008*a*, 2008*b*; Hansen 2006; Gjuvsland et al. 2011; Milocco and Salazar-Ciudad 2020). Some of these authors show that genetic covariances can change under a nonlinear GPM and that this change is often not what we would expect for a linear GPM (Houle 1991; Rice 2004). However, the mode and tempo of these changes for realistic GPMs is still not known.

In this study, we perform an exhaustive exploration of the evolution of *G* under the complex and nonlinear GPMs found in the study of development. For that purpose, we combine a computational GPM model that is based on our current understanding of the development of a complex organ, the mammalian tooth (Salazar-Ciudad and Jernvall 2010), and a population model with mutation, recombination, and directional selection. In brief, we have a population of genotypes and simulate the development of each individual in the population using the developmental model (see figs. 1, S1; figs. S1–S11 are available online). We then apply selection on the three-dimensional morphology resulting from the development of each individual. We also apply mutation and recombination in each generation. We estimate *G* in each generation to measure how it changes over time.

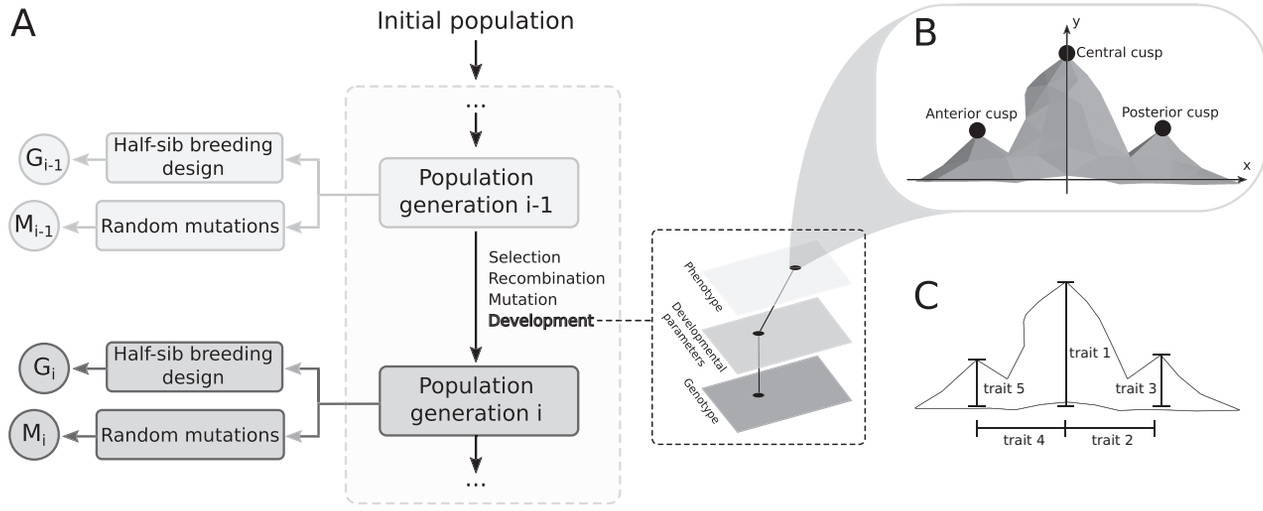

Figure 1: *A* shows the algorithm of the evolutionary simulations. Each individual in the population is modeled explicitly and has a set of genetic values, which additively determine the values of the developmental parameters. These parameters are mapped to a phenotype using our development model. Selection, recombination, and mutation are applied in each generation. G_i and M_i, the *G* and *M* matrices at generation *i*, respectively, are estimated with high precision in each generation (see "Methods"). *B* shows an example of a phenotype produced by the developmental model and the location of the three landmarks on the three tallest cusps of the tooth. *C* shows the traits measured and the coordinates of the landmarks in each individual's phenotype.

The mathematical model of tooth developmental is well suited for our objective because it is one of the few development models that has been able to reproduce multivariate morphological variation in a natural population (Salazar-Ciudad and Jernvall 2010) and has been previously used to simulate microevolution (Salazar-Ciudad and Marin-Riera 2013; Milocco and Salazar-Ciudad 2020). A scheme of the model is shown in figure S1. The model explicitly simulates a sheet of epithelial cells that fold and divide to produce a tooth morphology, recapitulating the process of tooth development. The model incorporates both the gene network known to be involved in tooth development and the mechanical interactions and cellular behaviors required for the formation of teeth. Mathematically, the model is a set of differential equations that describe how gene expression, cell position, and mechanical properties change in time for each cell during development. The initial condition of the tooth model is a small sheet of epithelial cells expressing a set of extracellular diffusible gene products. As developmental time proceeds, these gene products regulate the expression of other gene products, as well as cell behaviors like cell division and adhesion. As a result, the epithelium folds and tooth morphology changes. These changes in the morphology of the tissue in turn affect the spatial distribution of the gene products and consequently result in further changes in gene expression, cell behaviors, and ultimately how the tooth continues to fold. The result of the developmental simulation is a complex three-dimensional morphology, described as the specific distribution of cells in space.

The dynamics of the tooth development model depend on a set of developmental parameters. There are 21 of these parameters, and they specify different aspects like diffusion rates and cell mechanical properties (for details, see "Methods"; Salazar-Ciudad and Jernvall 2010). The values of these parameters are determined by the genotype of each individual, so genetic variation results in phenotypic variation by affecting the values of these parameters. However, the correspondence between developmental parameters and genes is largely unknown, and we do not explicitly simulate such relationship per se. For example, cell adhesion depends on the expression of multiple adhesion and cytoskeletal proteins (e.g., Jernvall and Thesleff 2012).

Interactions between gene products and cell behaviors as represented in the tooth model are a common feature of mathematical models of development (Oster and Alberch 1982; Raspopovic et al. 2014; Osterfield et al. 2017; Glen et al. 2019). On the basis of these models and what is known about organ development, it has been suggested that the dynamics of interaction among gene products, cell behaviors, and morphology is, overall, of a similar complexity in most organs and body parts and that, thus, the overall complexity and properties of their GPM should be similar (Alberch 1982; Salazar-Ciudad et al. 2003; Urdy 2012). Hence, although the GPM of the tooth model may differ from those of other systems in many aspects, its overall complexity is informative of that of several other systems, at least compared with a linear GPM. This is especially the case for morphological phenotypes, particularly for those that, like teeth, form by the folding of epithelia. This includes organs like heads, early brain, kidneys, genitalia, limbs, lungs, and insect wings (Gilbert and Barresi 2016).

In our population model, each individual has a genotype, a set of values of the developmental parameters, a phenotype, and a fitness. In each individual morphology we measure five traits describing tooth shape, shown in figure 1. Individual fitness is a function of the distance between the values of the traits of the individual and a set of optimal values. A different optimal morphology is used in each evolutionary simulation. We explore all possible combinations of upward and downward selection per trait, leading to $2^5 = 32$ optima. The distance to the optimum is used to determine which individuals are selected as parents for the next generation. Mutation is applied to individuals in each generation. The processes of development, selection, and mutation are iterated over generations to simulate evolution.

Evolution in our simulation can be seen as occurring in three different spaces: genotype space, developmental parameter space, and phenotypic space (see fig. 1). The genotype of an individual can be represented as a point in genetic space. This point is mapped linearly to a point in developmental parameter space, since developmental parameters are additively determined by genes. In turn, this point in developmental parameter space is mapped to a point in phenotype space through the developmental model. Therefore, the nonlinearity in the GPM arises only from the developmental model. Importantly, for different regions in developmental parameter space, the mapping to trait space can be different depending on the dynamics of the tooth model. In other words, the same change in the value of a developmental parameter will result in different magnitudes and directions of phenotypic changes, depending on the region of developmental parameter space in which an individual lies. The development model is deterministic and directly represents the GPM, without environmental effects.

For each generation, we estimate the G matrix by performing an in silico parallel half-sib breeding design (Lynch and Walsh 1998). We then measure how G changes over generations. We use most of the summary statistics that have been proposed in the literature to compare G matrices. These include the size of G (i.e., the sum of the eigenvalues of G; Jones et al. 2003; McGuigan 2006; Hansen and Houle 2008), the eccentricity of G (i.e., the proportion between

the largest eigenvalue and the sum of the rest; Jones et al. 2003, 2012; Kirkpatrick 2009), the change in direction of the main eigenvector of *G* (Eroukhmanoff and Svensson 2011; Johansson et al. 2012), and the amount of additive genetic variance in the direction of selection (Walsh and Blows 2009). We also use a method that directly compares *G* matrices as the difference in the predicted responses using the breeder's equation. The combined study of how these matrix properties change gives a rich picture of *G* matrix evolution, including different aspects in which *G* can change. Because we compare pairs of *G* matrices, we do not use other methods designed to compare several matrices, such as the tensor method (Hine et al. 2009). In addition to our in-depth study of the evolution of *G*, we also study how the *M* matrix evolves and how it is related to *G*. The *M* matrix describes the variance and covariance between traits due to the mutations entering the population in each generation, as opposed to *G*, which measures standing additive genetic variation.

Methods

Evolutionary Simulations

Details of the simulations are given in a previous article (Milocco and Salazar-Ciudad 2020). As briefly summarized below, we ran a total of 32 evolutionary simulations of 50 generations each using different optima to explore the behavior of the system. The *G* matrix was estimated in every generation. All populations were composed of 300 males and 300 females.

Genotypes are diploid with a total of 105 (haploid) loci. The 21 developmental parameters are each additively determined by five loci (i.e., the additive effects of 10 alleles are added together to determine the value of each developmental parameter). Each locus contributes to only one developmental parameter. The phenotype of each individual is obtained by running the tooth model on its developmental parameters. For details on the tooth developmental model, see the publication where it was introduced (Salazar-Ciudad and Jernvall 2010). In brief, the model explicitly simulates a sheet of epithelial cells, together with the gene network known to be involved in tooth development. Signaling (developmental) parameters numerically represent how growth factors diffuse and affect each other's expression. Growth factor also influence tissue dynamics by determining the rate of cell proliferation. The resulting shape of the tooth is additionally affected by other mechanical (developmental) parameters, such as adhesion and repulsion. A scheme of the model is given in figure S1. The output of the model is a tooth morphology. We measure five traits, namely, the coordinates of three landmarks located on the three tallest cusps of the tooth (see fig. 1).

Directional selection is applied on the population by selecting the 150 males and 150 females (i.e., 50% of the population) that are closest to an optimum morphology defined at the beginning of each evolutionary simulation. Selection is applied in this way for the 50 generations, resulting in larger or smaller selection differentials depending on the available variation at a given time. Each evolutionary simulation has a different optimum morphology (see "Supplementary Methods" in the supplemental PDF, available online). All possible combinations of upward and downward selection for the five traits are explored, leading to $2^5 = 32$ optima. We only study directional selection, as the populations do not reach the optimum within the 50 generations of the simulations.

Selected parents are paired randomly. Parents contribute one of their two alleles in each loci to the offspring with equal probability. There is a 10^{-3} chance that a loci is mutated in the process. Mutation is implemented by adding a random number to the value of an allele, drawn from a normal distribution with a zero mean and a standard deviation equal to 0.1. Each parental couple produces two female and two male offspring. This means that all selected individuals contribute the same number of offspring to the next generation, while unselected individuals do not contribute. Once the genotypes of the offspring are determined, the cycle of generating each individual's morphology, measuring the traits, selection, and mutation, is iterated for the next generation. The initial population was generated by running the same iteration of steps but with stabilizing selection for 400 generations.

Estimation of the G Matrix

First the evolutionary simulation was run for the 50 generations. Then the population at generation i ($i = 1, 2, \ldots, 50$) was copied and used as the parental population of a half-sib breeding design to estimate G_i, the *G* matrix in generation *i*. The breeding design was made without selection (i.e., random sampling of parental population) following four steps: (1) sample half of the males from the parental population, (2) sample half of the females from the parental population, (3) form couples randomly, and (4) generate four offspring per couple. Steps 1 through 4 were repeated a total of 10 times, generating a data set composed of full sibs and half-sibs. This data set is used for a single estimate of G_i (see fig. 1) using an animal model. Restricted maximum likelihood estimates of G_i were obtained using the software WOMBAT (Meyer 2007). The animal model used was the simplest possible, $y_j = \mu + a_j + e_j$, where y_j is the phenotype of individual j, μ is the population mean, a_j is the additive genetic merit of individual j, and e_j is a random residual error.

Estimation of the M Matrix

The M matrix for generation i was estimated by making a copy of all individuals of generation i and introducing single-locus mutations to all loci of all individuals (see fig. 1). We then ran the tooth development model on the one-locus mutants, obtaining a mutant phenotype. We calculated the effect of the mutation as the difference between the mutant phenotype and the original phenotype. The M matrix is then obtained as the covariance matrix of these mutational effects.

Measures of Change in G

As explained in the introduction, there are several ways to measure how G changes. In this work, we use many of these measures, each capturing different aspects of change in G. In what follows, we compare G_i and G_{i-k}, where G_i is the G matrix at generation i and k is the number of generations separating the two compared G matrices, what we call the "gap." We refer to λ_j as the jth eigenvalue of G, from largest to smallest (i.e., λ_1 is the largest eigenvalue). We note that matrix comparisons in this work are done directly in the scale that the traits are measured, since all traits considered are distances of the same order of magnitude (see fig. 1). The measurements of change in G used in this work are listed below.

Change in response to selection. This measure is defined as $\|G_{i-k}\beta_{i-k} - G_i\beta_{i-k}\|/\|G_{i-k}\beta_{i-k}\|$, where $\|.\|$ is the Euclidean vector norm and β_i is β at generation i. This is an intuitive measure that describes the difference between G_i and G_{i-k} as the relative difference in the responses to selection they predict when used in the breeder's equation, together with the β acting on the population at time $i - k$.

Change in size of G. Here, size is defined as $\sum_j \lambda_j$ (Steppan et al. 2002; Arnold et al. 2008). This is a measure of the total amount of additive genetic variance in G.

Change in direction of G. Here, the direction is defined by $g_{\max}$, the eigenvector of G with the largest eigenvalue. This is the direction in trait space where there is the most additive genetic variance. We calculate the change in $g_{\max}$ as the angle between $g_{\max_i}$ and $g_{\max_{i-k}}$, where $g_{\max_i}$ is the $g_{\max}$ of G_i.

Change in effective dimension of G. This measure is defined as $\sum_j \lambda_j/\lambda_1$. It was proposed by Kirkpatrick (2009) and considers how evenly distributed the genetic variance is over the eigenvectors of G.

Change in eccentricity of G. This measure is defined as $\lambda_1/\sum_j \lambda_j$. This is the inverse of the effective dimension and describes how much of the total additive genetic variance is distributed along the main eigenvector.

Change in the amount of additive genetic variance in the direction of β. This measure is calculated as $\beta^{\mathrm{T}}G\beta$, with the superscript T the transpose operator.

Element-by-element change. This measure is defined as the absolute value of the difference between the corresponding elements of the compared G matrices.

Results

How Much Does G Change?

We found that the G matrix changes during our simulations of phenotypic evolution. Figure 2*A* shows the change in G for an example simulation (see fig. S2 for the rest of simulations). In this figure, the change in the G matrix between generation i and generation $i - 1$ was measured as the relative difference in the predicted response to selection using the two corresponding G matrices in the multivariate breeder's equation, with the same selection gradient (see "Methods"). For generation i, this was calculated as $\|G_{i-1}\beta_{i-1} - G_i\beta_{i-1}\|$, normalized by $\|G_{i-1}\beta_{i-1}\|$, where G_i and β_i are the G matrix and the selection gradient at generation i and $\|.\|$ is the Euclidean norm (i.e., the vector length). We call this measure the "change in G" in figure 2, and it provides an intuitive measure of the change in the G matrix as how much that change affects the predicted response to selection. We find that this change can be large. In the simulation shown in figure 2*A*, change is relatively large between generations 10 and 30. This corresponds to a period when the phenotypes in the population change dramatically, as seen from the morphologies included in the figure. Figure S3 shows the change for each element of G for this simulation, with the largest changes measured for the variance of trait 2 and its associated covariances. We also measured change as the relative difference in the size of G and the angle change in $g_{\max}$ (i.e., the direction of greatest additive genetic variance) between consecutive generations. These other measurements of change (fig. S4) show behavior similar to the one based on the breeder's equations.

The changes in G over generations can have different magnitudes, and even in a single generation these changes can be large. Figure 2*B* shows the distribution of change in G between consecutive generations for all of the generations of all of our 32 evolutionary simulations. The median relative change in G is 5% between consecutive generations, and 10% of these single-generation changes are larger than 24%.

We compared G matrices at different generation intervals. We call this generation interval the "gap" (i.e., for a gap of k, we compare G_i and G_{i-k}). Figure 2*C* shows the distribution of change in G as a function of the gap. As expected, when comparing G matrices separated by a larger gap, the measured change is larger, since G has more time to evolve. However, even for small gaps, G can change appreciably. In some exceptional situations, large changes

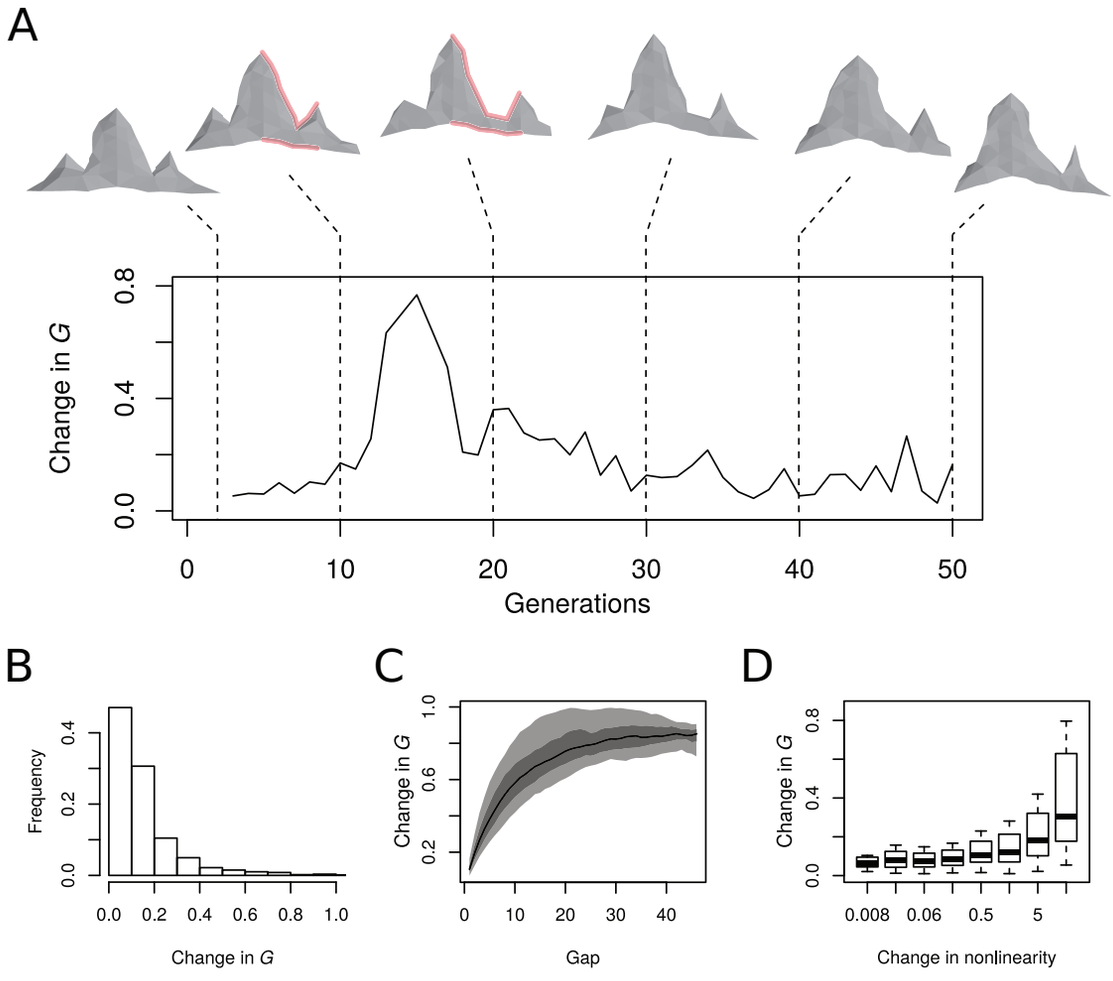

Figure 2: *G* changes in time. The changes can be large, and they are related to the genotype-phenotype map (GPM). We compare the response to selection predicted using G_i and G_{i-k}, with k being the generation gap. *A* shows the change with a gap of 1 for an example simulation (simulation 6; all simulations are shown in fig. S1). The teeth represent the morphologies closest to the population mean at different generations. In red we show the part of the morphology that undergoes a relatively large change. This change is associated with the relatively large changes in *G* in those generations. *B* shows the distribution of the change in *G* for a gap of 1, for all generations of all simulations. *C* shows the distribution of the change in *G* as a function of the gap for all generations and all simulations. The black line is the median change for each gap, and the light gray area and the dark gray area contain 40% and 20% of the distribution of change, respectively. Note that the variance of the distribution tapers off for large gap sizes, since there are fewer comparisons that can be made for larger gaps. *D* shows that there is a relationship between the change in nonlinearity of the GPM (see "Supplementary Methods" in the supplemental PDF) and the magnitude of the change in *G*.

in *G* of even 100% can occur in only a few generations (see, e.g., simulation 18 in fig. S2). Figure S5 shows the distribution of change in *G* against the gap for other measures of change in *G*—namely, size and angle of g_{max}—displaying distributions similar to the one shown in figure 2*C*.

The amount of change is related to the change in the nonlinearity of the GPM, as shown in figure 2*D*. The most dramatic changes in *G* occur when the nonlinearity of the GPM changes more rapidly. We measure the nonlinearity of a region of the GPM as how well the map is locally approximated by a linear fit (i.e., the nonlinearity of the GPM is the norm of the residuals of the best linear fit: if it is small, it means that the GPM is locally described by a linear map; see "Supplementary Methods" in the supplemental PDF). We then calculate the change in this measurement between consecutive generations, capturing how the local characteristics of the GPM are changing. Both the median and the spread of the changes in *G* increase with the change in nonlinearity. The increase in the median change in *G* indicates that larger changes in *G* can occur when the nonlinearity changes more rapidly. The increase in the spread of change in *G* indicates

that both large and small changes occur when nonlinearity changes rapidly. The same relationship with the nonlinearity of the GPM is seen with other measures of change in *G*, as shown in figure S5.

Classifying the Change in G

The *G* matrix changes in several qualitative different ways. These can be studied by following the behavior of three measurements (see "Methods"). The first measurement is the size of *G*, which is the sum of the eigenvalues of *G* and describes the total amount of additive genetic variance. The second measurement is the direction of g_{max}, which describes the orientation of the main axis of variation. We measure it as the angles of g_{max} with each of the vectors of an arbitrary orthogonal basis of five-dimensional space. The third measurement is the eccentricity of *G*, which is the leading eigenvalue divided by the sum of eigenvalues. This describes how much of the total additive genetic variance is distributed along the main axis of variation.

The eccentricity and size of *G* are functions of its eigenvalues, shown for all simulations against the generations in figure S6. Using these measurements, we define five categories for the change in *G*. Category A is constant *G*, characterized by constant eccentricity and size. Category B is proportional change in *G*, where the size of *G* changes while keeping the same eccentricity. This means that *G* either shrinks of expands. Category C is disproportional change in *G*, where the eccentricity of *G* changes and the size may or may not vary. Category D is rotation of *G*, where the direction of g_{max} changes. Category E is sudden changes in *G*, which are relatively rapid changes in eccentricity, size, and/or orientation of *G* that can occur in nonlinear regions of the GPM. Note that categories A, B, C, and E are mutually exclusive. Category D can be combined with A, B, or C.

Figure 3*A* shows a descriptive scheme of the categories; examples for each of the categories taken from the simulations are shown in figure S7. The *G* matrices are plotted as 95% confidence ellipses with the direction and length of the axes determined by the two eigenvectors with the largest eigenvalues (see "Supplementary Methods" in the supplemental PDF). This classification resembles the Flury hierarchy for comparing *G* matrices (Steppan et al. 2002; Arnold et al. 2008). However, we do not use hypothesis testing in our classification since we have very precise estimates of

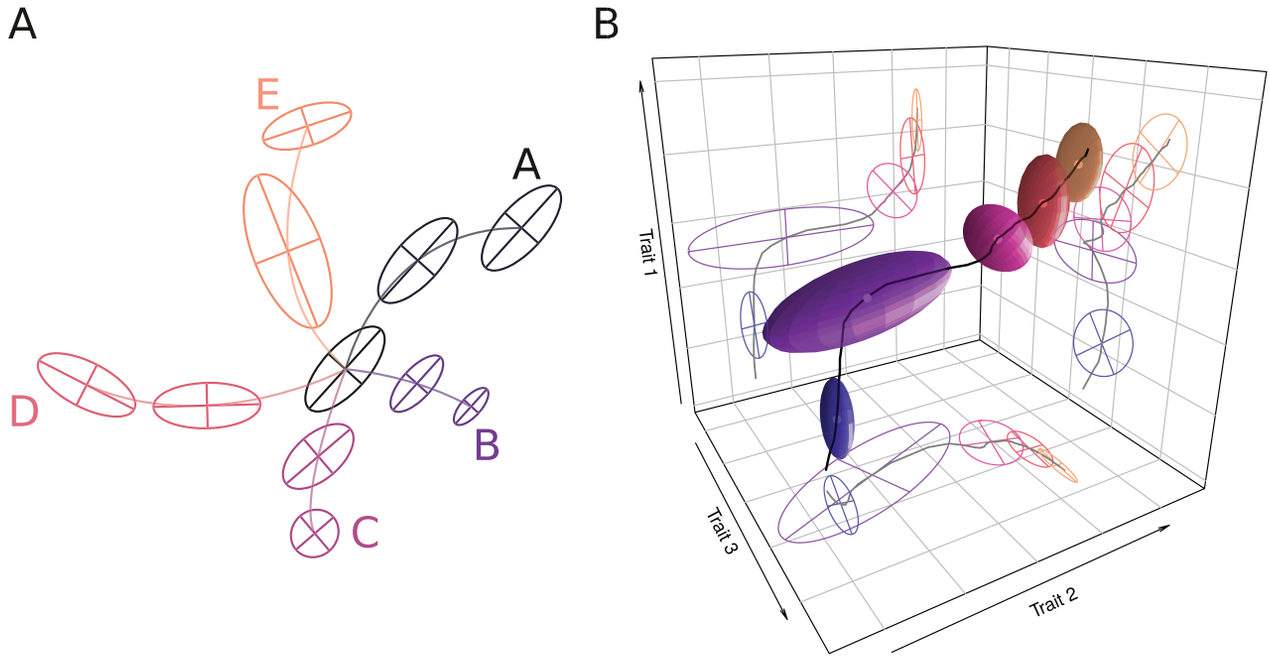

Figure 3: *G* can change in several ways. *A* shows a schematic representation of the types of changes that can be observed in *G* in our simulations. Figure S7 shows examples of these dynamics found in the simulations. *B* shows the evolution of the *G* matrix in an example simulation. The plot shows the *G* matrix for five generations of simulation 6, also shown in figure 2. For this plot, we use the 3 × 3 submatrix of *G* corresponding to traits 1, 2, and 3 (see fig. 1). We plot this submatrix for generations 7 (blue), 17 (purple), 27 (pink), 37 (red), and 47 (yellow). Each ellipsoid is also projected in the two-dimensional planes, where ellipses represent the 2 × 2 submatrix of *G* corresponding to a pair of traits (i.e., traits 1 and 2, traits 2 and 3, or traits 1 and 3). The plot shows that in a single simulation *G* can change dramatically and that it can exhibit several different types of changes described in the main text.

G due to a lack of environmental effects and very large sample sizes.

Using the size, eccentricity, and angle of g_{max} in each generation, we perform an automatic classification of the dynamics of *G* (see "Supplementary Methods" in the supplemental PDF). For the classification, we compare *G* matrices separated by a generation gap of 3, 6, and 9 because single-generation comparisons are typically associated with changes that are too small for reliable classifications. We define thresholds for each measurement: if the difference in a measurement between compared *G* matrices is larger than the threshold, we consider that aspect of *G* to have changed. Table S1 (available online) shows the frequency of each of the categories, using all data from all simulations and different thresholds. Category A (i.e., constant *G*) is the most common in all gaps and thresholds considered. This is expected, since we are not considering large gaps. Importantly, all other dynamics appear in significant proportions as well.

Several of the categories of changes in *G* can occur in a single simulation run. Figure 3*B* shows an example simulation where the *G* matrices for three of the five traits are plotted as ellipsoids in time. The two-dimensional projections of these ellipsoids are plotted on the side panels, describing the *G* matrices for pairs of traits. During this simulation, *G* changes rapidly in a nonlinear region (category E), rotates (category D), and becomes more eccentric (category C) at different generations as it evolves.

Our measure of eccentricity is the inverse of the effective dimension of *G* (Kirkpratrick 2009). The dimension of *G* can be seen as the subspace of the phenotypic space in which evolutionary change is possible. In most empirical studies, the observed dimension is smaller than the number of traits measured (Kirkpatrick and Lofsvold 1992; Hine and Blows 2006; McGuigan and Blows 2007; but see Mezey and Houle 2005). We found that in our simulations, the effective dimension of *G* can change (see fig. 4 for an example simulation and fig. S8 for the rest of simulations). To better visualize the change in dimension (fig. 4*B*), we projected the *G* matrices in the three different generations marked with arrows in figure 4*A* into a three-dimensional space. The projection captures the overall shape of *G* at all times, with more than 92% of variance explained by the axes (see "Supplementary Methods" in the supplemental PDF). At generation 2 *G* is quite flat (i.e., effectively two-dimensional), but it becomes completely three-dimensional around generation 15. Then *G* again becomes a (two-dimensional) disc by generation 23.

Additive Genetic Variance in the Direction of Selection

Figure 5*A* shows the evolution of additive genetic variance in the direction of selection for simulation 10. Additive genetic variance (V_A) can increase, decrease, or remain constant during the response to directional selection (other simulations are shown in fig. S9). In the simulation shown in figure 5, V_A decreases from generation 2 to generation 15 and then remains constant until generation 30, when it starts to increase. The dynamics are highly dependent on the local characteristics of the GPM, as shown by the representations of the GPM included in the figure. Indeed, the increase in V_A after generation 30 occurs because the population enters a nonlinear region of the GPM where much adaptive variation is suddenly available, leading to an increase in V_A in the direction of selection. Note, however, that the amount of nonadditive genetic variance in those regions increases proportionally more.

Evolution of the M Matrix

The *M* matrix is the variance-covariance matrix of mutational effect and summarizes the variation that enters the population in a generation through mutations. Figure 5*A* shows that mutational variance (V_M) in the direction of selection can increase, decrease, or stay constant depending on the local GPM, for an example simulation (the rest of the simulations are shown in fig. S9). We find that *G* and *M* are similar in terms of the distribution of variance and in the angle between the main directions of variation (i.e., their first eigenvector), g_{max} and m_{max}, respectively (fig. S10).

Temporal misalignment of g_{max} and m_{max} can occur when the population crosses a nonlinear region of the GPM (fig. 5*C*). The *M* matrix describes variation entering the population at a given generation, which immediately changes when the population enters a nonlinear region of the GPM. The *G* matrix, on the other hand, summarizes standing genetic variation and therefore lags behind the changes in *M*. That lag results in temporal misalignments of g_{max} and m_{max}. The matrices realign because new mutations become incorporated into standing genetic variation.

Discussion

G Can Change Rapidly

Our results show that the *G* matrix can evolve rapidly. We simultaneously compared several measures of change in *G*, giving a rich picture of the different aspects of *G* matrix structure that can change. These measures include changes in the total amount of additive genetic variance (the size of *G*), changes in the distribution of variation along axes (eccentricity and effective dimension), changes the projection of *G* in the direction of selection, changes in the direction of the axes of variation (orientation of g_{max}), and differences in the response to selection predicted from *G* matrices

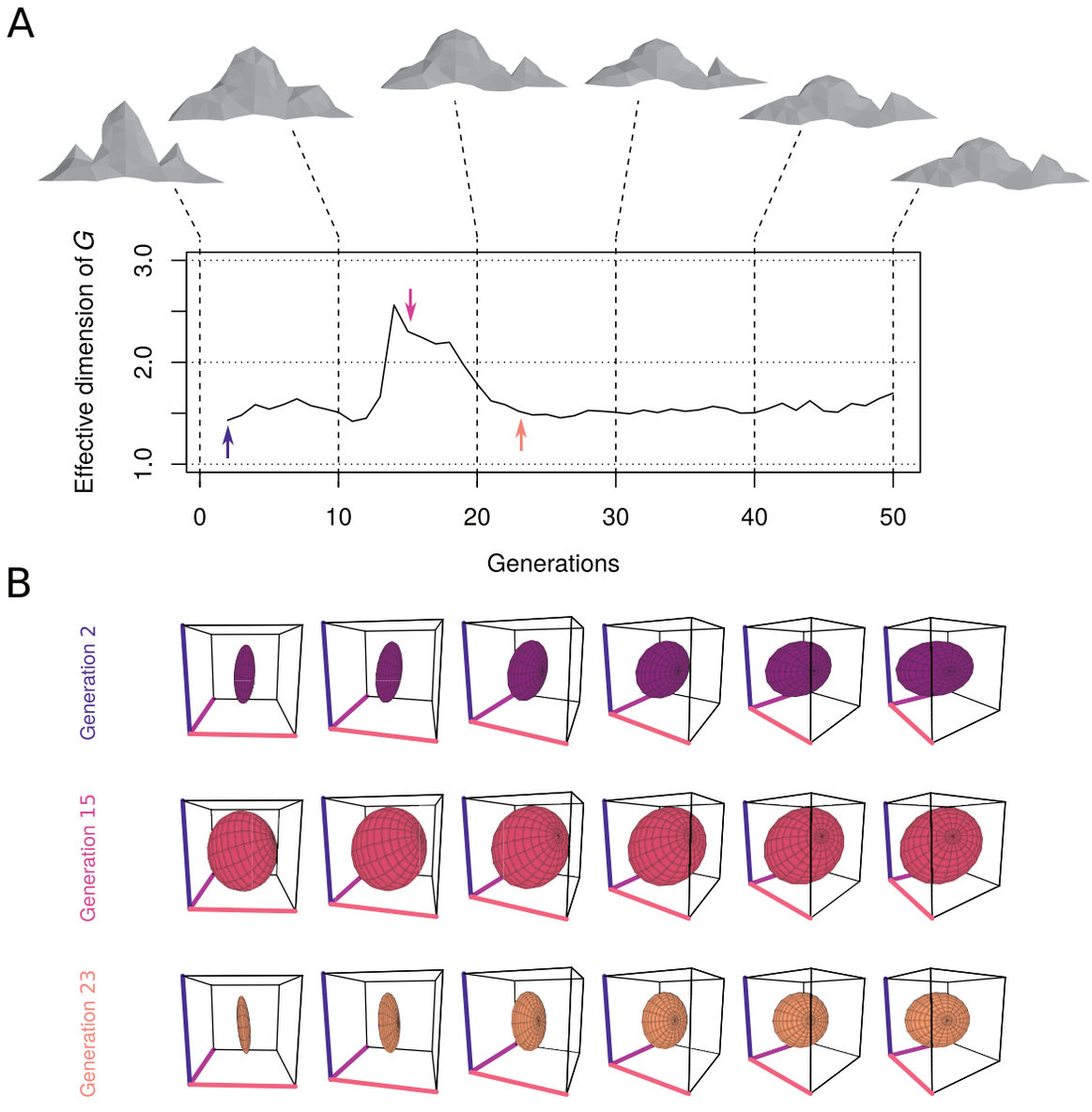

Figure 4: The effective dimension of G changes in our simulation. *A* shows the effective dimension for each generation of example simulation 18 (other simulations are shown in fig. S8). Dotted lines are included at dimensions 1, 2, and 3 for reference. Between generations 10 and 20, G goes from having an effective dimension between 1 and 2 to having one between 2 and 3. The teeth closest to the population mean at different generations are included. *B* shows the projection of G in three-dimensional space (see "Supplementary Methods" in the supplemental PDF), at the three different generations marked in *A* with arrows. The projections conserve most of the additive variance (92%, 98%, and 94% for generations 2, 15, and 23, respectively). The same ellipsoid representing G in three-dimensional space is shown using slightly different angles for each generation.

estimated at different generations. For all of these measurements we found that G can change quickly, with changes of more than 100% in less than 10 generations. However, we also found situations where all aspects of G remained unchanged for several generations.

We have found that the evolution of G is determined by an interplay between the GPM and selection. Selection pushes the population to move in the trait space toward the optimal phenotype. Moving in the trait space implies moving across the developmental parameter and genotypic space too (see fig. 6). The dynamics of development are different for different values of the developmental parameters, leading to different morphologies with different trait values and a different relationship to genetic variation.

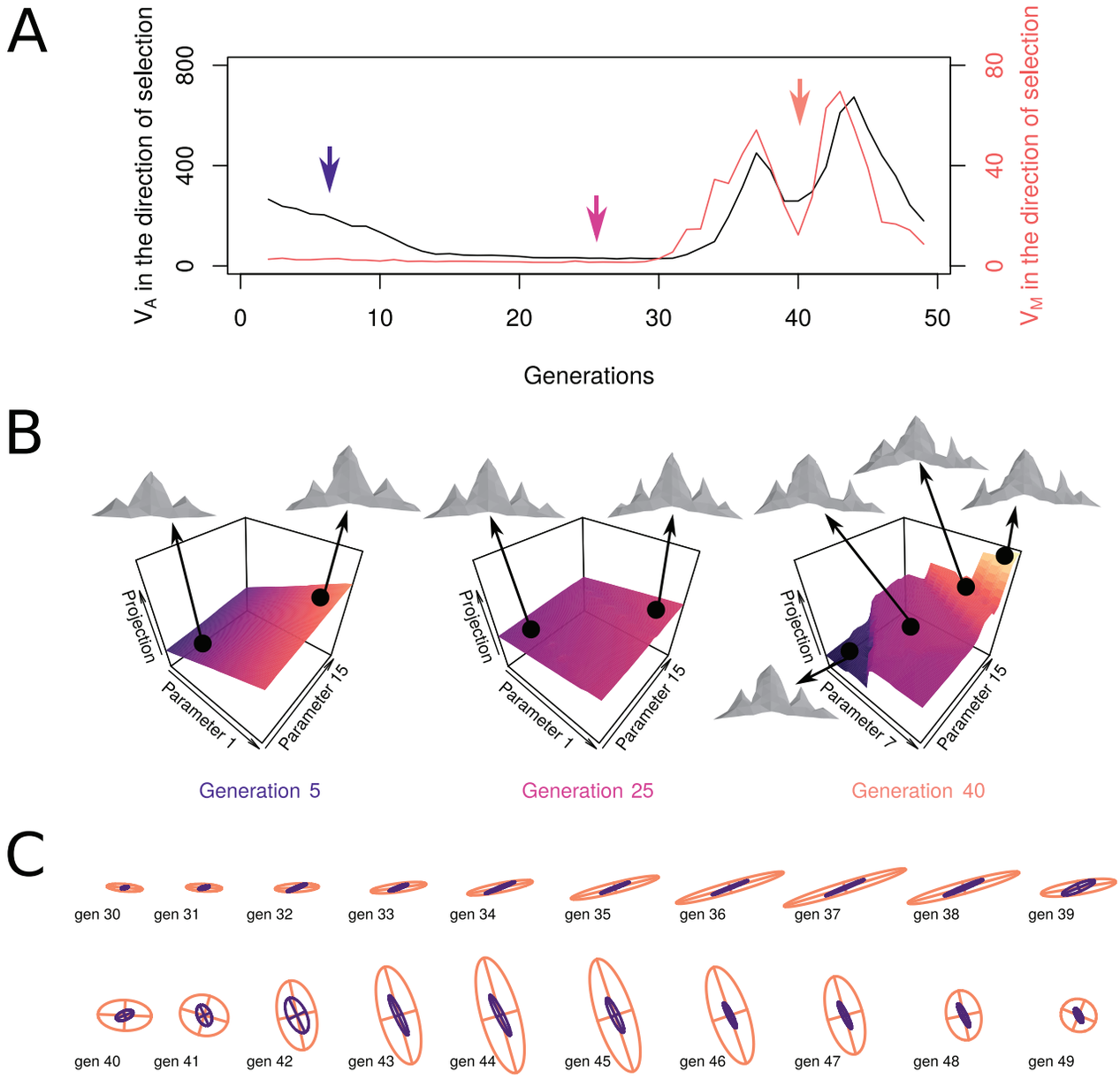

Figure 5: Evolution of *M* and *G*. *A* shows the evolution of additive genetic variance (V_A) and mutational variance (V_M) in the direction of selection for simulation 10 (other simulations are shown in fig. S9). These amounts can increase, decrease, or remain constant during the response to directional selection. The dynamics are highly dependent on the local characteristics of the genotype-phenotype map (GPM), as can be seen from the marginal GPMs at the three time points shown in *B*. Marginal GPMs were constructed by tinkering with the two developmental parameters (*xy*-plane) that explain most phenotypic variance around their mean value. The *z*-axis in these maps is the deviation of the tinkered phenotypes away from the mean phenotype, projected in the direction of selection (see "Supplementary Methods" in the supplemental PDF). *C* shows projections of *M* and *G* in two-dimensional space for the same simulation. The two matrices are largely aligned, with *G* lagging behind *M*.

For example, as shown in figure 6, a small change in the value of a parameter can lead to a relatively large change in the traits in some regions of the parameter space but not in others. As the genetic makeup of the population changes under selection, the population moves across the developmental parameter space, and there can be a change in the local GPM it experiences. Different local GPMs lead to different morphological variation and covariation (see fig. S11) and therefore to different *G* matrices. If the local GPM changes rapidly, so will the *G* matrix. In turn, if

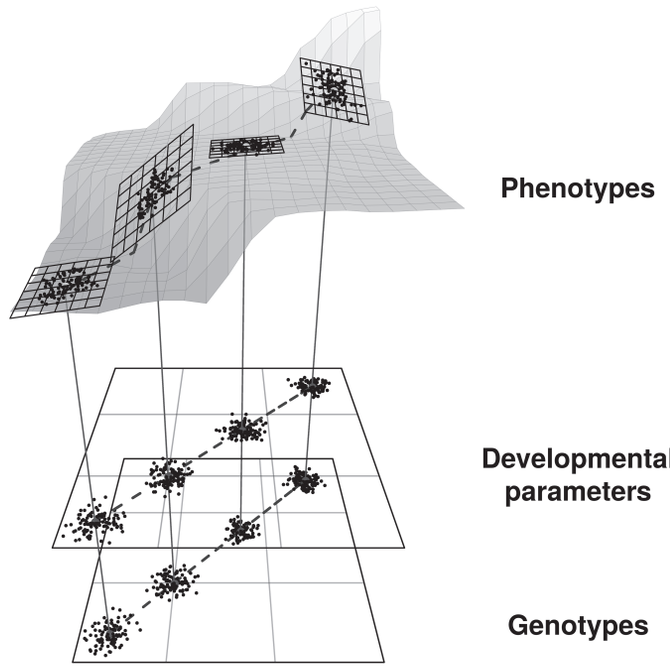

Figure 6: The local characteristics of the genotype-phenotype map (GPM) change as the population evolves. The population at four different generations is plotted as a cloud of points. For every generation, each point on the genetic space represents the genetic values of one individual. Genetic values additively determine the developmental parameters, so each individual is mapped linearly to developmental parameter space. The developmental parameters determine the phenotype through the dynamics of the development model. This results in the mapping from developmental parameters to phenotype to be locally different as the population evolves. We include linear approximations of the GPM at the different generations, represented as a tangent plane plotted in black. This is the linear approximation of the GPM that is assumed in *G* matrix models (see Rice 2008*b*). The orientation of this plane changes dramatically as the population evolves. The same change in genetic values results in a different phenotypic change in different regions of the GPM. These different regions of the GPM also result in different ways in which traits covary, as shown in figure S11. Note that this figure is a simplification. As the population evolves, all 21 developmental parameters change, not only two, as plotted here.

the local GPM stays the same while the population is evolving, *G* will remain constant.

Previous Theoretical Approaches: Linear GPM Models Do Not Show the Same Diversity of Changes in G as Seen in Our Model

Previous studies of the evolution of *G* use GPM models with assumptions that are quite different from ours (Turelli 1985; Slatkin and Frank 1990; Reeve 2000; Agrawal 2001; Jones et al. 2003, 2012, 2014). The large majority of theoretical work on the evolution of *G* assumes a linear GPM, where each allele is modeled as having a direct additive genetic effect on traits (i.e., each allele increases or decreases a trait in a given amount that is not affected by other alleles). Trait values are then modeled as being determined by the sum of the additive effects of alleles on traits. Jones and collaborators have a series of seminal articles studying *G* evolution on a linear GPM using individual-based simulations (Jones et al. 2003, 2012; Arnold et al. 2008). More recently, the authors expanded their investigation including epistatic interactions (Jones et al. 2014) using the multilinear model (Hansen and Wagner 2001), which assumes a linear influence of the genetic background on the effect of an allele. The general expectation coming from this body of work using linear GPMs is that the expected shape of *G* is ultimately determined by selection, with random fluctuations around this expected shape (Jones et al. 2003, 2012, 2014).

In our model, phenotypes are not defined as the sum genetic effects. In contrast, phenotypes are determined by networks of interactions between gene products, the regulation of cell behaviors, and mechanical properties involved in the development of a complex organ. Mathematically, this model consists of a set of coupled differential equations, one per cell, that define how gene expression, cell position, and mechanical properties change in time for each cell. The equations have a set of developmental parameters that in this study are additively determined by alleles. The equations are nonlinear, and the total number of them varies during simulation time since there is cell division. Furthermore, none of these equations is directly

assigned to the traits that we measure. Cells push each other in complex patterns over developmental time. Thus, we do not know a priori which cells will be in the tallest cusps, which define the traits we measure. Furthermore, the fact that cells push each other during development implies that all equations (i.e., all cells) can have an effect on trait values, even if such an effect is indirect and complex. In other words, none of our equations prespecifies the phenotype. Phenotypes and, thus, the values of the traits emerge from the dynamics of the model.

As a result of affecting a developmental parameter, an allele can have an effect on phenotypic variation, but this effect varies depending on the values of the other developmental parameters. In other words, alleles do not have intrinsic or fixed phenotypic effects. Thus, we do not impose the distribution of allelic effects, as other models of the evolution of *G* usually do. On the contrary, this distribution emerges from the development model. Even if one can statistically estimate the additive effect of an allele in a population, such an effect can be defined only for a given generation and population (i.e., a given region of the developmental parameter space), since there is no guarantee that it would not change, even dramatically, in one or a few generations. Similarly, the development model does not specify epistasis or pleiotropy among alleles. Instead, both epistasis and pleiotropy arise from the dynamics of the developmental model and can vary for different regions of the parameter space.

The nature of the GPM we use in our simulations results in changes in *G* that can be larger than those reported under the assumption of a linear GPM. We also find a large diversity of types of changes that cannot be found in simulations using a linear GPM but have been found in nature and experimentally (see "Comparison with Empirical Measures of *G* Evolution" below). Most importantly, as further explained below, the evolution of *G* in our model depends not only on selection but also on the GPM the population experiences as it evolves.

Should G Align with Natural Selection?

An aspect of *G* and *M* that has received special attention is their projection on the direction of selection. This is the amount of additive genetic (V_A) and mutational (V_M) variance in the direction of selection, respectively. When properly standardized, these amounts are called evolvability (Hansen and Houle 2008). Some authors suggest that these quantities should increase in evolution as a direct result of selection (Pavlicev 2011; Jones et al. 2012, 2014). This would occur by selection increasing the frequency of the alleles associated with favorable correlations among traits and decreasing the frequency of those associated with unfavorable correlations. As a result, *G* and *M* would reorient to increase their projection in the direction of selection (Pavlicev 2011; Jones et al. 2014). This assumes that there can be genetic variation for any direction of change in *G*—in other words, that mutations can change *G* in any conceivable direction. However, realistically, there are some rules on how the GPM—and as a result, *G*—can change. As in the case of morphological variation (Alberch 1982), not all directions of variation are necessarily possible or equally likely for a given *G*. Each region of the developmental parameter space can lead to different morphologies, a different GPM, a different *G*, and different types of change in *G*.

In our case, the rules of change in *G* come from the tooth development model itself. An alignment among *G*, *M*, and the direction of selection is likely to occur only in models where there are no specific rules of change or evolution in the GPM. In this scenario, selection is able to shape the structure of variation without restrictions. In contrast, in our model an alignment among *G*, *M*, and selection is not to be expected in general. The population simply moves in the developmental parameter space to reach more adaptive phenotypes, and this will result in different *G* matrices. These matrices will reflect the local characteristics of the GPM where the population is at that point in time (fig. 6). Here, it is important to remember that the properties of the GPM depend on the tooth model and are not themselves affected by natural selection (at least on the timescale we consider; see "Caveats" below). Natural selection does affect what regions of the GPM the population is crossing, but only indirectly through selection on the phenotype.

In fact, we have found that V_A and V_M in the direction of selection can increase, decrease, or remain constant depending on the trajectory of a population across the GPM in our simulations (see fig. 5). This has been previously suggested for the single-trait case and in specific terms of the directionality of epistasis (Carter et al. 2005; Hansen 2006).

That V_A can sometimes increase during the evolutionary simulations, even in the direction of selection, may seem counterintuitive. Natural selection leads to the fixing of alleles that move the phenotype toward the optimum and, thus, decrease phenotypic and genetic variation in this direction (Zhang and Hill 2005). This is certainly the case, but as allelic frequencies change the population also crosses different regions of the GPM. As explained above, these regions have different properties (i.e., a different local GPM) that lead to different *G* matrices and, in turn, changes V_A.

In some cases, V_A can even increase quickly for a small number of generations and decrease shortly after (see fig. 5*A* for an example). This occurs when the population is crossing a region of the developmental parameter space

where further change of a developmental parameter leads to a relatively large change, or a "jump," in one or several traits. Note that this relatively large change is still gradual to the eye, as shown with the morphologies in figure 5*B*, for example. When the population is in this region, some of the individuals will be distributed before the jump, and others will be distributed after the jump. As a result, both V_A and V_P peak. In fact, V_P increases more than V_A because there is also an increase in nonadditive genetic variance. Once the whole population has crossed the jump, however, both V_A and V_P will decrease.

Comparison with Empirical Measures of G Evolution

There is a large body of empirical work studying the evolution of *G* (reviewed in Steppan et al. 2002; Arnold et al. 2008; Wood and Brodie 2015). Empirical studies of the evolution of *G* are typically of one of two types. The first type comprises studies comparing *G* matrices among closely related natural populations that have experienced different environments for a certain amount of time. These studies are very diverse in the timescales considered (from tens to thousands of generations separating the compared populations). The second type comprises studies in laboratory populations under controlled conditions, typically with the objective of assessing the impact of specific experimental treatments on *G*. Our results are most directly comparable to the results from this second type of study, since our simulations are for 50 generations with persistent directional selection and other factors controlled, similar to an experimental setup. Unlike empirical work, however, we are able to estimate *G* with high precision in every generation, and in that sense our results cannot be directly compared to existing data. Indeed, because of the logistic difficulty in estimating *G*, most empirical studies include only two estimates of *G*.

There is empirical evidence for all of the types of changes in *G* we found in our study. Category A (no change in *G*) was clearly shown, for example, by Hangartner et al. (2020) for four traits in *Drosophila melanogaster* populations that have locally adapted along a latitudinal cline over approximately 100 years. Category A was also found by Delahaie et al. (2017), who studied changes in *G* for seven traits among four populations of blue tits (*Cyanistes caeruleus*).

Category B (proportional changes in *G*) has been reported in Blows and Higgie (2003) for eight traits in an experimental treatment with *Drosophila serrata* for nine generations. Category C (nonproportional changes in *G*) was reported by Doroszuk et al. (2008) for three life history traits of the soil nematode *Acrobeloides nanuswas* in response to benign and stress conditions for 20 years. The authors additionally found evidence of category D (rotation in *G*) in their data. Categories C and D were also found together in a study by Walter et al. (2018) of an Australian native wildflower, *Senecio lautus*, for 10 traits, among populations that have diverged less than half a million years ago and that occupy four distinct habitats.

Eroukhmanoff and Svensson (2011) studied the change in *G* matrix between two ecotype populations (ancestral and derived) of an aquatic isopod, *Asellus aquaticus*. The authors measured seven traits and found evidence for changes in the *G* matrix in categories C and D. The authors also found evidence of change in the dimension of *G*. Specifically, they found a reduction in the dimension of *G* in the recently derived ecotype.

Rapid changes in *G* (category E in this work) have been reported in natural settings by Björklund et al. (2013). They found rapid and large changes in *G* in a 25-year-long study of the collared flycatcher (*Ficedula albicollis*) for four morphological traits. They found large changes in size, shape, and orientation of the estimated *G* matrices. The authors suggest that the temporal changes in *G* could be explainable by environment-by-genotype effects, which have been found to have the potential to alter genetic covariances as fast as in a single generation (Sgró and Hoffmann 2004; Wood and Brodie 2015; Sniegula et al. 2018).

An important clarification is that in our simulations we can estimate the *G* matrix with high precision in each generation, since there are no environmental effects in our model, sample sizes are large, and we have negligible measurement error. Empirical studies instead are associated with much larger sample variances. This is particularly the case for studies in the wild, where the availability of data is more restricted. This means that part of the changes in the *G* matrix from empirical studies could be explained by sampling variance.

Caveats

We do not include environmental effects in our simulations. Morphology arises during development, and the environment can affect such morphology by affecting the developmental dynamics. Environmental effects in our model can be introduced, for example, as noise added to the developmental parameters before generating the phenotype using the development model. In a previous study using the same simulations of evolution (Milocco and Salazar-Ciudad 2020), we show that the inclusion of environmental effects does not change the dynamics of evolution in our simulations. Environmental effects do not change the GPM, but they can affect the regions of the developmental parameter space over which the individuals in a population are spread. A possible future study of environmental effects could model environmental effects in

a more realistic way than noise. Interestingly, that would result in a mapping from genotype-by-environment space to developmental parameters.

Another caveat of our work is that the GPM model we use is based on a tooth development model. Other systems develop in different ways and, hence, result in GPMs that are different from the one in teeth. In that sense, our results should be regarded as indicative of the general properties that many GPMs may have or, at least, as an alternative to a linear view of the GPM. By "general properties" we mean that there will be regions of the GPM that are flat, while others will show large nonlinearities. For the focus of this article, which is the evolution of *G*, this means that for the GPM of other organs we can also expect *G* to evolve differently in these different regions—show abrupt changes in some regions, remain constant in others, show changes in dimension, and so on.

A further limitation of our approach is that we consider genetic variation only within the genetic and cellular interactions included in our model (i.e., the developmental model itself does not change). The reason for this choice is that our study is focused on a relatively short timescale, on the order of 10^1–10^2 generations. On this timescale, we can expect that changes in the structure of the developmental process are unlikely. This is because changing the development process would imply acquiring new gene interactions or the regulation of additional cell behaviors (e.g., a change in the topology or logic of the underlying developmental process). On longer timescales, however, we expect our results to hold. This is because we expect that for complex phenotypes such as teeth, the GPMs are unlikely to evolve to be more linear. This expectation stems from two lines of reasoning, explained below.

First, morphologies have to be physically constructed during development (e.g., from a zygote) through complex networks of gene and cell interactions and a limited set of cell behaviors (e.g., cell division, cell contraction, cell adhesion, cell secretion; Alberch 1982; Salazar-Ciudad et al. 2003; Urdy 2012). Many of these interactions are intrinsically nonlinear (Alberch 1982; Newman and Müller 2000), and thus development is an intrinsically nonlinear process.

Second, theoretical work supports the view that the when development evolves to be able to produce complex morphologies it leads, as a side effect, to quite complex GPMs (Newman and Müller 2000; Salazar-Ciudad et al. 2001; Hagolani et al. 2021). Related theoretical work suggests that GPMs can become more linear over evolutionary time but that this is very unlikely for complex morphologies (Salazar-Ciudad et al. 2001; Salazar-Ciudad and Jernvall 2004). Essentially building a complex morphology that has a simple GPM, when possible, requires a very complex development (with many genes and gene interactions), while complex morphologies with a complex GPM can be built by simpler developmental networks. We then expect complex morphologies with simple GPMs to evolve rarely because they require the accumulation of many mutational changes.

Conclusions

Our results indicate that complex GPMs lead to different qualitative and quantitative theoretical expectations of how *G* evolves compared with those of linear GPMs. Selection moves the population through different regions of the GPM. The different regions of the GPM that the populations explore have different characteristics, resulting in different phenotypic variation being produced as well as a different local GPM and therefore determining how selection can further proceed. This dialectical interaction between selection and the GPM determines how the *G* matrix evolves.

In agreement with other researchers (e.g., Eroukhmanoff 2009; Björklund et al. 2013), our results suggest that one should be very careful when using the *G* estimated in a generation to infer evolution at later generations (or for related populations). In fact, contrary to what previous theoretical studies suggest, our expectation should be that *G* can change relatively fast and thus that these kind of inferences are unlikely to be reliable. Our study also provides a detailed description of the qualitatively different ways in which *G* could evolve. Most of these types of change have not been previously reported in theoretical work. Thus, by using a more realistic GPM model, we can obtain a richer and more detailed view of the rates and modes of *G* evolution.

Acknowledgments

We thank Pascal Hagolani, Miguel Brun-Usan, Hugo Cano Fernández, Carlos Mora, Renske Vroomans, and Aleksa Ratarac for useful comments. This research was funded by the Center of Excellence in Experimental and Computational Developmental Biology, the Finnish Academy (project 315740), the Spanish Ministry of Science (PGC2018-096802-B-I00), and the Doctoral Programme in Integrative Life Science of the University of Helsinki.

Statement of Authorship

L.M. and I.S.-C. collaborated in the conceptualization, funding acquisition, and writing (review and editing). L.M. wrote the original draft and performed the data analysis and visualization. I.S.-C. supervised the study.

Data and Code Availability

The *G* matrices for all generations and simulations used in this work have been deposited in the Dryad Digital Repository (https://doi.org/10.5061/dryad.z34tmpgck; Milocco 2021).

Literature Cited

Agrawal, A. F., E. D. Brodie, and L. H. Rieseberg. 2001. Possible consequences of genes of major effect: transient changes in the **G**-matrix. Genetica 112/113:33–43.

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.

Alberch, P. 1982. Developmental constraints in evolutionary processes. Pages 313–332 *in* Evolution and development. Springer, Berlin.

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

Björklund, M., A. Husby, and L. Gustafsson. 2013. Rapid and unpredictable changes of the **G**-matrix in a natural bird population over 25 years. Journal of Evolutionary Biology 26:1–13.

Blows, M. W., and M. Higgie. 2003. Genetic constraints on the evolution of mate recognition under natural selection. American Naturalist 161:240–253.

Cano, J. M., A. Laurila, J. Palo, and J. Merilä. 2004. Population differentiation in **G** matrix structure due to natural selection in *Rana temporaria*. Evolution 58:2013–2020.

Careau, V., M. E. Wolak, P. A. Carter, and T. Garland. 2015. Evolution of the additive genetic variance–covariance matrix under continuous directional selection on a complex behavioural phenotype. Proceedings of the Royal Society B 282:20151119.

Carter, A. J. R., J. Hermisson, and T. F. Hansen. 2005. The role of epistatic gene interactions in the response to selection and the evolution of evolvability. Theoretical Population Biology 68:179–196.

Chakrabarty, A., and H. Schielzeth. 2020. Comparative analysis of the multivariate genetic architecture of morphological traits in three species of Gomphocerine grasshoppers. Heredity 124:367–382.

Delahaie, B., A. Charmantier, S. Chantepie, D. Garant, M. Porlier, and C. Teplitsky. 2017. Conserved **G**-matrices of morphological and life-history traits among continental and island blue tit populations. Heredity 119:76–87.

Doroszuk, A., M. W. Wojewodzic, G. Gort, and J. E. Kammenga. 2008. Rapid divergence of genetic variance-covariance matrix within a natural population. American Naturalist 171:291–304.

Eroukhmanoff, F. 2009. Just how much is the **G**-matrix actually constraining adaptation? Evolutionary Biology 36:323–326.

Eroukhmanoff, F., and E. I. Svensson. 2011. Evolution and stability of the **G**-matrix during the colonization of a novel environment. Journal of Evolutionary Biology 24:1363–1373.

Falconer, D. S., and T. F. C. Mackay. 1996. Introduction to quantitative genetics. 4th ed. Pearson, London.

Gilbert, S. F., and M. J. F. Barresi. 2016. Developmental biology. Sinauer, Sunderland, MA.

Gjuvsland, A. B., J. O. Vik, J. A. Woolliams, and S. W. Omholt. 2011. Order-preserving principles underlying genotype-phenotype maps ensure high additive proportions of genetic variance. Journal of Evolutionary Biology 24:2269–2279.

Glen, C. M., M. L. Kemp, and E. O. Voit. 2019. Agent-based modeling of morphogenetic systems: advantages and challenges. PLoS Computational Biology 15:e1006577.

Hagolani, P. F., R. Zimm, R. Vroomans, and I. Salazar-Ciudad. 2021. On the evolution and development of morphological complexity: a view from gene regulatory networks. PLoS Computational Biology 17:e1008570.

Hangartner, S., C. Lasne, C. M. Sgrò, T. Connallon, and K. Monro. 2020. Genetic covariances promote climatic adaptation in Australian *Drosophila*. Evolution 74:326–337.

Hansen, T. F. 2006. The evolution of genetic architecture. Annual Review of Ecology, Evolution, and Systematics 37:123–157.

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

Hansen, T. F., and G. P. Wagner. 2001. Modeling genetic architecture: a multilinear theory of gene interaction. Theoretical Population Biology 59:61–86.

Hine, E., and M. W. Blows. 2006. Determining the effective dimension of the genetic variance-covariance matrix. Genetics 173:1135–1144.

Hine, E., S. F. Chenoweth, H. D. Rundle, and M. W. Blows. 2009. Characterizing the evolution of genetic variance using genetic covariance tensors. Philosophical Transactions of the Royal Society B 364:1567–1578.

Houle, D. 1991. Genetic covariance of fitness correlates: what genetic correlations are made of and why it matters. Evolution 45:630–648.

Jernvall, J., and I. Thesleff. 2012. Tooth shape formation and tooth renewal: evolving with the same signals. Development 139:3487–3497.

Johansson, F., M. I. Lind, P. K. Ingvarsson, and F. Bokma. 2012. Evolution of the **G**-matrix in life history traits in the common frog during a recent colonisation of an island system. Evolutionary Ecology 26:863–878.

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

Jones, A. G., R. Bürger, and S. J. Arnold. 2014. Epistasis and natural selection shape the mutational architecture of complex traits. Nature communications 5:1–10.

Jones, A. G., R. Bürger, S. J. Arnold, P. A. Hohenlohe, and J. C. Uyeda. 2012. The effects of stochastic and episodic movement of the optimum on the evolution of the **G**-matrix and the response of the trait mean to selection. Journal of Evolutionary Biology 25:2210–2231.

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

Kirkpatrick, M., and D. Lofsvold. 1992. Measuring selection and constraint in the evolution of growth. Evolution 46:954–971.

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

Lynch, M., and B. Walsh. 1998. Genetics and analysis of quantitative traits. Sinauer, Sunderland, MA.

McGuigan, K. 2006. Studying phenotypic evolution using multivariate quantitative genetics. Molecular Ecology 15:883–896.

McGuigan, K., and M. W. Blows. 2007. The phenotypic and genetic covariance structure of drosphilid wings. Evolution 61:902–911.

Merilä, J., M. Björklund, and L. Gustafsson. 1994. Evolution of morphological differences with moderate genetic correlations among traits as exemplified by two flycatcher species (Ficedula;

Muscicapidae). Biological Journal of the Linnean Society 52:19–30.

Meyer, K. 2007. WOMBAT: a tool for mixed model analyses in quantitative genetics by restricted maximum likelihood (REML). Journal of Zhejiang University Science B 8:815–821.

Mezey, J. G., and D. Houle. 2005. The dimension of genetic variation for wing shape in *Drosophila melanogaster*. Evolution 59:1027–1038.

Milocco, L. 2021. Data from: Evolution of the *G* matrix under nonlinear genotype-phenotype maps. American Naturalist, Dryad Digital Repository, https://doi.org/10.5061/dryad.z34tmpgck.

Milocco, L., and I. Salazar-Ciudad. 2020. Is evolution predictable? quantitative genetics under complex genotype-phenotype maps. Evolution 74:230–244.

Müller, G. B. 2007. Evo-devo: extending the evolutionary synthesis. Nature Reviews Genetics 8:943–949.

Newman, S. A., and G. B. Müller. 2000. Epigenetic mechanisms of character origination. Journal of Experimental Zoology 288:304–317.

Oster, G., and P. Alberch. 1982. Evolution and bifurcation of developmental programs. Evolution 36:444–459.

Osterfield, M., C. A. Berg, and S. Y. Shvartsman. 2017. Epithelial patterning, morphogenesis, and evolution: *Drosophila* eggshell as a model. Developmental Cell 41:337–348.

Pavlicev, M., J. M. Cheverud, and G. P. Wagner. 2011. Evolution of adaptive phenotypic variation patterns by direct selection for evolvability. Proceedings of the Royal Society B 278:1903–1912.

Penna, A., D. Melo, S. Bernardi, M. I. Oyarzabal, and G. Marroig. 2017. The evolution of phenotypic integration: how directional selection reshapes covariation in mice. Evolution 71:2370–2380.

Phelan, J. P., M. A. Archer, K. A. Beckman, A. K. Chippindale, T. J. Nusbaum, and M. R. Rose. 2003. Breakdown in correlations during laboratory evolution. I. Comparative analyses of *Drosophila* populations. Evolution 57:527–535.

Raff, R. A. 1996. The shape of life: genes, development, and the evolution of animal form. University of Chicago Press, Chicago.

Raspopovic, J., L. Marcon, L. Russo, and J. Sharpe. 2014. Digit patterning is controlled by a Bmp-Sox9-Wnt Turing network modulated by morphogen gradients. Science 345:566–570.

Reeve, J. P. 2000. Predicting long-term response to selection. Genetics Research 75:83–94.

Rice, S. H. 2004. Developmental associations between traits: covariance and beyond. Genetics 166:513–526.

———. 2008*a*. Theoretical approaches to the evolution of development and genetic architecture. Annals of the New York Academy of Sciences 1133:67–86.

———. 2008*b*. The G-matrix as one piece of the phenotypic evolution puzzle. Evolutionary Biology 35:106–107.

Salazar-Ciudad, I., and J. Jernvall. 2004. How different types of pattern formation mechanisms affect the evolution of form and development. Evolution and Development 6:6–16.

———. 2010. A computational model of teeth and the developmental origins of morphological variation. Nature 464:583–586.

Salazar-Ciudad, I., J. Jernvall, and S. A. Newman. 2003. Mechanisms of pattern formation in development and evolution. Development 130:2027–2037.

Salazar-Ciudad, I., S. A. Newman, and R. V. Solé. 2001. Phenotypic and dynamical transitions in model genetic networks. I. Emergence of patterns and genotype-phenotype relationships. Evolution and Development 3:84–94.

Salazar-Ciudad, I., and M. Marín-Riera. 2013. Adaptive dynamics under development-based genotype-phenotype maps. Nature 497:361–364.

Sgrò, C. M., and A. A. Hoffmann. 2004. Genetic correlations, tradeoffs and environmental variation. Heredity 93:241–248.

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.

Slatkin, M., and S. A. Frank. 1990. The quantitative genetic consequences of pleiotropy under stabilizing and directional selection. Genetics 125:207–213.

Sniegula, S., M. J. Golab, S. M. Drobniak, and F. Johansson. 2018. The genetic variance but not the genetic covariance of life-history traits changes towards the north in a time-constrained insect. Journal of Evolutionary Biology 31:853–865.

Steppan, S. J., P. C. Phillips, and D. Houle. 2002. Comparative quantitative genetics: evolution of the **G** matrix. Trends in Ecology and Evolution 17:320–327.

Turelli, M. 1985. Effects of pleiotropy on predictions concerning mutation-selection balance for polygenic traits. Genetics 111:165–195.

Urdy, S. 2012. On the evolution of morphogenetic models: mechano-chemical interactions and an integrated view of cell differentiation, growth, pattern formation and morphogenesis. Biological Reviews of the Cambridge Philosophical Society 87:786–803.

Walsh, B., and M. W. Blows. 2009. Abundant genetic variation + strong selection = multivariate genetic constraints: a geometric view of adaptation. Annual Review of Ecology, Evolution, and Systematics 40:41–59.

Walter, G. M., J. D. Aguirre, M. W. Blows, and D. Ortiz-Barrientos. 2018. Evolution of genetic variance during adaptive radiation. American Naturalist 191:E108–E128.

Wood, C. W., and E. D. Brodie III. 2015. Environmental effects on the structure of the **G**-matrix. Evolution 69:2927–2940.

Zhang, X. S., and W. G. Hill. 2005. Genetic variability under mutation selection balance. Trends in Ecology and Evolution 20:468–470.

Associate Editor: Stephen F. Chenoweth
Editor: Jennifer A. Lau