Evolution of adaptive phenotypic variation patterns by direct selection for evolvability

Mihaela Pavlicev1,*,†, James M. Cheverud1 and Günter P. Wagner2

1Department of Anatomy and Neurobiology, Washington University, 660 South Euclid Avenue, PO Box 8108, St Louis, MO 63110, USA
2Department of Ecology and Evolutionary Biology, Yale University, 165 Prospect Street, PO Box 208106, New Haven, CT 06520-8106, USA

A basic assumption of the Darwinian theory of evolution is that heritable variation arises randomly. In this context, randomness means that mutations arise irrespective of the current adaptive needs imposed by the environment. It is broadly accepted, however, that phenotypic variation is not uniformly distributed among phenotypic traits, some traits tend to covary, while others vary independently, and again others barely vary at all. Furthermore, it is well established that patterns of trait variation differ among species. Specifically, traits that serve different functions tend to be less correlated, as for instance forelimbs and hindlimbs in bats and humans, compared with the limbs of quadrupedal mammals. Recently, a novel class of genetic elements has been identified in mouse gene-mapping studies that modify correlations among quantitative traits. These loci are called relationship loci, or relationship Quantitative Trait Loci (rQTL), and affect trait correlations by changing the expression of the existing genetic variation through gene interaction. Here, we present a population genetic model of how natural selection acts on rQTL. Contrary to the usual neo-Darwinian theory, in this model, new heritable phenotypic variation is produced along the selected dimension in response to directional selection. The results predict that selection on rQTL leads to higher correlations among traits that are simultaneously under directional selection. On the other hand, traits that are not simultaneously under directional selection are predicted to evolve lower correlations. These results and the previously demonstrated existence of rQTL variation, show a mechanism by which natural selection can directly enhance the evolvability of complex organisms along lines of adaptive change.

Keywords: modularity; allometry; canalization; differential epistasis

1. INTRODUCTION

It has long been recognized that phenotypic variation is highly structured [1–4]. The best documented pattern is called morphological integration as identified by Olson & Miller [5], where functionally related traits tend to be more strongly correlated than functionally less-related traits. Morphological integration has been documented in many systems, from plants to vertebrate animals (e.g. [6–10]) and also confirmed at the level of quantitative trait locus (QTL) effects rather than phenotypic or genetic correlations (e.g. [11–14]). Furthermore, it has been shown that patterns of integration change in evolution in a predictable way. For example, in quadrupedal animals, the corresponding elements in the forelimbs and hindlimbs are more correlated than in animals with specialized limbs. For instance, in bats, the forelimbs and hindlimbs perform different functions and have lower correlations than those of other mammals [15]. Moreover, the lower correlation among forelimbs and hindlimbs in hominoid primates is associated with a greater disparity in forelimb and hindlimb proportions than in monkeys [16]. All these data suggest that correlation patterns among traits evolve in a predictable way and influence the evolvability of the affected traits. These observations raise the question of what the evolutionary mechanisms that shape the pattern of phenotypic variation are. How much is genetic variation affecting the covariation among phenotypic characters? How does natural selection mould these variation patterns? The first question has been answered by QTL mapping and mutation accumulation studies which show that there is ample genetic variation affecting the covariation among characters (see below). In this paper we address the second question by modelling how selection acts on that form of genetic variation.

A variety of methods have been used to determine the amount of genetic variation affecting trait correlations. For instance, a recent quantitative genetic study compared correlations between traits that vary across families in a population [17]. The results show that a substantial proportion (about 40%) of this variation is genetic. Evidence for such variation has also been found in previous experimental work. For example, Lenski [18,19] documented genetic variation in Escherichia coli of the association between resistances to two different phages. Templeton and colleagues [20,21] found that the expression and association of the name-giving phenotype abnormal abdomen with other phenotypic traits that are affected by the same locus, varies across

* Author for correspondence (mihaela.pavlicev@bio.uio.no).
† Present address: Department of Biology, Center for Ecological and Evolutionary Synthesis, University of Oslo, PO Box 1066, Blindern, 0316 Oslo, Norway.

Electronic supplementary material is available at http://dx.doi.org/10.1098/rspb.2010.2113 or via http://rspb.royalsocietypublishing.org.
fly populations. Recently, Kim et al. [22] documented genetic variation in correlation of phenotypic traits in yeast, causing varying degrees of rust coloration being associated with different degrees of drug resistance.

To map the loci associated with genetic variation in trait correlation, Cheverud et al. [23] developed a QTL mapping method. By associating marker genotypes with the regression coefficient between two phenotypic traits, it was found that there are many polymorphic loci that affect the statistical relationships among phenotypic traits ([17,23,24]; figure 1b). These loci have been called ‘relationship Quantitative Trait Loci’ (rQTL), because they change the variational relationships between characters. Genotypes can differ in the degree of canalization of variance both owing to gene-by-gene and gene-by-environment interaction. Two recent papers ([17], see also [24]) established that the effect of the rQTL is because of differential epistatic interactions, i.e. epistatic interactions are trait-specific. Epistasis was also found responsible for variable intertrait associations in studies of single loci [18–22]. In brief, the change in genetic background affects the traits of a pleiotropic domain of a locus in different ways, leading to differences in correlations among traits across genotypes at the focal locus (figure 1b). When multiple traits are affected by epistasis simultaneously, differential canalization of traits changes intertrait correlations through differentially modifying the expression of pre-existing genetic variation. Interestingly, rQTL mapping studies have shown that a substantial proportion of these rQTL have no detectable additive effect. These ‘neutral’ QTL (i.e. neutral with respect to the trait mean) constitute 30–84% of the rQTL loci detected in mice (75% in Cheverud et al. [23]; 30% in Pavlicev et al. [17]; 84% in Leamy et al. [24]). This fraction of ‘neutral’ rQTL is particularly interesting because selection can act on their variational effects without conflicting with selection on their additive effects. Here, we present a phenomenological model of selection on these neutral rQTL, i.e. genes that change the variance and covariance among traits but do not directly affect the mean phenotype (see scheme in figure 1b). We show that this model can lead to the alignment of high levels of phenotypic variation with the direction of selection (figure 1c).

2. THE MODEL

To investigate how directional selection affects rQTL variation, we construct a simple haploid model of a population with two alternative genotypes at an rQT locus. We model the life cycle in three steps: in the first step, we model the evolution of the quantitative traits within each rQTL genotype class using Lande’s phenotypic models [25,26]. Because each rQTL allele causes a different genetic covariance matrix, the selection response is different in the two rQTL genotype classes. This will lead to different average fitness of the two rQTL genotypes. In the second step, we model selection among the two rQTL classes using a deterministic one-locus, two-allele model. Finally, we model sexual reproduction as an incomplete mixture process between the genotypic values associated with each rQTL allele.

Specifically we consider a phenotype with two traits \( z_1 \), \( z_2 \), with sexual reproduction and a haploid genotype.

Modelling the haploid case aids transparency. However, the same scenario can be modelled for a diploid genome, although with considerably greater bookkeeping costs. The rQTL has two alternative alleles, \( A_i \) and \( A_j \). We assume random mating, absence of gene-by-environment interaction and a constant \( G \) within each rQTL genotype class. The two alternative genotypes at the haploid rQTL differ in genotypic (\( G_j \) and \( G_j \)) and phenotypic (\( P_i \) and \( P_j \)) covariance matrices, but initially not in their average phenotypic values (figure 1b). The mean fitness values of the two rQTL genotype classes are denoted as \( w_1 \), \( w_2 \), and \( w_1 \), \( w_2 \).
the difference between them determines the selection coefficient of the rQTL alleles, $A_i$ and $A_j$.

The genotypic values of the next generation are then derived in three steps. In the first step, the trait means of the selected population are derived according to Lande’s $\Delta \bar{z} = G \nabla \ln \bar{w}$. The model is based on the selection regime (see below) and the respective genetic covariance matrices, with

$$\nabla \ln \bar{w} = \left( \frac{\partial \ln \bar{w}}{\partial z_1}, \ldots, \frac{\partial \ln \bar{w}}{\partial z_n} \right).$$

In our case with two traits, we write for rQTL allele $A_i$:

$$\Delta \bar{z}_i = G \nabla \ln \bar{w},$$

$$\Delta \bar{z}_i = \left[ \bar{g}_{11,i} \bar{g}_{12,i} \bar{g}_{21,i} \bar{g}_{22,i} \right] \left[ \frac{\partial \ln \bar{w}}{\partial z_1} \ldots \frac{\partial \ln \bar{w}}{\partial z_2} \right].$$

In most of our scenarios we assume constant directional selection on both traits, with the fitness function $w(z_1, z_2) = w_0 \exp[a_1 z_1 + a_2 z_2]$, where $w_0$ is an arbitrary scaling coefficient [27]. The coefficients $a_1$ and $a_2$ determine the strength of directional selection on traits $z_1$ and $z_2$, respectively. We use the multiplicative form of selection because a linear selection function causes a decrease in selection gradient with increasing population mean, while the selection gradient remains constant with multiplicative fitness. Using the exponential fitness function $w(z) = w_0 \exp[a_1 z_1 + a_2 z_2]$ and the bivariate normal distribution for the distribution of phenotypic values, the mean fitness reads as

$$\bar{w} = w_0 \exp[a_1 \bar{z}_1 + a_2 \bar{z}_2 + M],$$

with

$$M = \frac{1}{2} \left[ a_1^2 V_1 + 2a_1 a_2 V_{12} + a_2^2 V_2 \right].$$

Differentiating the log mean fitness with respect to mean values gives

$$\frac{\partial \ln \bar{w}}{\partial z_1} = a_1 \quad \text{and} \quad \frac{\partial \ln \bar{w}}{\partial z_2} = a_2,$$

respectively.

Thus, following the selection of the breeding population, the change in traits affected by the two genotypes is

$$\Delta \bar{z}_{11} = a_1 \bar{g}_{11} + a_2 \bar{g}_{12},$$

$$\Delta \bar{z}_{21} = a_1 \bar{g}_{12} + a_2 \bar{g}_{22},$$

$$\Delta \bar{z}_{1j} = a_1 \bar{g}_{1j} + a_2 \bar{g}_{12},$$

$$\Delta \bar{z}_{2j} = a_1 \bar{g}_{12} + a_2 \bar{g}_{2j}.$$

In the second step, the fitness differences between the rQTL classes are used to determine the frequency change of the rQTL alleles according to Wright’s selection equation:

$$p' = \frac{\bar{w} p}{\bar{w}}$$

for $A_i$ with the frequency $p$, and analogously for $A_j$ with frequency $q$.

In the third step, the recombination between the rQTL and the genetic background affecting trait means is implemented as a mixture process of genotypes. This step reflects recombination during the transition to the next generation. Thus, the mean phenotypic value of the genotype $i$ in the next generation $(\bar{z}_i)$ depends on the phenotypic value of both genotypes in the previous generation $(\bar{z}_1^S, \bar{z}_2^S)$, to the extent of recombination:

$$\bar{z}'_i = (p + qL) \bar{z}_1^S + (q - qL) \bar{z}_2^S,$$

where $p, q$ are the allele frequencies of the two alleles at rQTL, and $L$ is a measure of linkage between the rQTL and other loci affecting the mean, $L \in [0, 0.75]$, with $L = 1$ representing full linkage and $L = 0.75$, free recombination. Linkage acts oppositely to recombination, maintaining the linkage disequilibrium (LD) between the advantageous alleles that build up owing to selection. Thus, instead of modelling recombination between pairs of loci explicitly, we model its composite effect on the phenotypic value of the particular rQTL genotype using the term $L$ (see appendix A for derivation and an alternative method). The rationale for the range of linkage values is the following: if $L$ would be 0.5, each mating among genotypes with different rQTL would lead to offspring with complete linkage equilibrium. Since free recombination at most leads to one-half of complete linkage equilibrium, the lowest level of linkage in this model is given by the mean between $L = 0.5$ and $L = 1$.

The resulting genotypic trait means of the next generation are used to calculate the corresponding mean fitness of the two rQTL genotypes, $\bar{w}_1, \bar{w}_2$, in the next generation:

$$\bar{w}_i = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} w(z_1, z_2)p(z_1, z_2|P_i) \ dz_1 dz_2.$$

Implementing recombination with the mean genotypic values generates trait means in the next generation. Owing to the different genetic covariances in the two rQTL genotype classes, the traits respond differently to selection within the rQTL classes. The genotype class with more genetic variation in the direction of selection will have a greater selection response and thus also acquire higher mean fitness than the alternative genotype class. Finally, the fitness differences between rQTL genotype classes determine the change in frequency of the rQTL alleles. The difference in mean fitness of the two rQTL genotype classes can be calculated as:

$$\Delta \bar{w}' = L \left[ \Delta \bar{w} + \left( a_1^2 \bar{g}_{11} + 2a_1 a_2 \bar{g}_{12} + a_2^2 \bar{g}_{22} \right) \right]$$

$$- \left( a_1^2 \bar{g}_{11} + 2a_1 a_2 \bar{g}_{12} + a_2^2 \bar{g}_{22} \right) \right],$$

where $\Delta \bar{w}$ is the fitness difference of rQTL genotypes in the previous generation, and $\bar{g}_{ij}$ are the elements of the genetic variance–covariance matrix for the two traits. The first two indices determine the coefficient in the $G$ matrix and $i$ or $j$ stand for either allele at the rQTL. We simplify equation (2.1) by replacing the constant terms with $S_i = a_1^2 \bar{g}_{11} + 2a_1 a_2 \bar{g}_{12} + a_2^2 \bar{g}_{22}$ and $S_j = a_1^2 \bar{g}_{11} + 2a_1 a_2 \bar{g}_{12} + a_2^2 \bar{g}_{22}$ and $S_i - S_j$ resulting in the pleasantly simple recursion equation

$$\Delta \bar{w}'(i,j) = L \left( \Delta \bar{w} + \Delta S \right).$$

(2.2)
When $\Delta \bar{w}(i,j) \neq 0$, selection occurs among the rQTL alleles. Because the genotypes at the rQTL are associated with different $G$ matrices, selection on the rQTL changes the covariance structure of the traits.

Several conclusions can be drawn from equation (2.2). First, we can see that the fitness difference between the two genotypes classes will increase or decrease depending on the sign of $\Delta S$. This term gives the selection response of the two genotypes classes to directional selection, depending on the rQTL allele-specific genetic covariance matrices and the direction and strength of selection. This change in mean fitness diminishes owing to recombination depending on the degree of linkage $L$. Note that even at the lowest level of linkage, the fitness difference does not completely disappear, because the LD does not relax within one generation.

Next, the bivariate genotypic values and the difference in mean fitness can be analytically derived in any generation. Given that the two alleles at rQTL have equal fitness in the first generation, iterating equation (2.2) over $N$ generations, $\Delta \bar{w}$ is predicted to be

$$\Delta \bar{w}^{(N)} = \frac{\Delta S(L^N + L^{N-1} + \cdots + L)}{L} = \frac{\Delta S L^N}{1 - L}.$$

Eventually, the mean fitness difference reaches a stable fixed point if linkage is not perfect, $L < 1$. The stationary fitness difference is

$$\Delta \bar{w} = \frac{L \Delta S}{1 - L}.$$

After the fitness difference has equilibrated, selection at the rQT locus proceeds with constant selection coefficients. This simplifies the analysis of selection dynamics.

We also explored the situation where one trait is under directional selection while the second trait is under stabilizing selection, a fitness function known as the ‘corridor model’ [28,29]. We applied an analogous logic as above but a different combination of selective pressures to select for dissociation of traits. This model is described by the following fitness function:

$$w(z_1, z_2) = w_0 \exp[a z_1 - b z_2^2],$$

where $a$ is a coefficient of directional selection acting on trait $z_1$, and $b$ is a coefficient of stabilizing selection acting on trait $z_2$.

The average fitness for the subpopulation with allele $i$ at the rQTL is then

$$\bar{w}_i = w_0 \sqrt{\frac{1}{1 + 2bg_{z_2}} \exp \left[ a^2 g_{z_1i} - 2az_1 - \frac{a^2 b g_{z_1i}^2 + a z_1 g_{z_1i} - b z_2^2}{1 + 2bg_{z_2}} \right]},$$

and differentiation yields the following selection gradients:

$$\frac{\partial \ln \bar{w}}{\partial z_1} = a \quad \text{and} \quad \frac{\partial \ln \bar{w}}{\partial z_2} = \frac{2b(a g_{z_2} - z_2)}{1 + 2bg_{z_2}}.$$
While this might be plausible in the case of long-term climatic trends or artificial selection, this assumption seems to be extreme. We therefore considered scenarios where directional selection is fluctuating, meaning that a pair of co-selected characters experience simultaneous selection for increasing or decreasing the trait means, as in situations where fluctuating temperatures may lead to alternating selection for greater or lesser body size. From the symmetry properties of equations (2.1) and (2.2), it can be seen that the same rQTL allele would be selected regardless of whether the traits are co-selected for larger or smaller means. It follows that fluctuating co-selection should also lead to increased absolute correlations among co-selected traits even when the direction of selection is fluctuating. We confirmed this intuition with simulations shown in figure 3.

The simulations presented in figure 3 are analogous to the previous ones in figure 2a,b, except that the signs of both selection coefficients change simultaneously every 100 generations. We have implemented two scenarios: in figure 3a, the population is allowed to reach linkage equilibrium between episodes of directional selection. This scenario simulates decay of positive linkage equilibrium during a short period of stabilizing selection between epochs of directional selection. In the second scenario in figure 3b, selection is reversed immediately following the previous epoch of directional selection.

Proc. R. Soc. B (2011)
The difference between these scenarios is important. LD develops between the rQTL alleles and the rest of the genes owing to selection. The rQTL allele causing faster evolution is associated with genes that cause trait means more advanced in the direction of selection. When the direction of selection changes, the advantageous rQTL allele is associated with the ‘wrong’ genes and is temporarily selected against until the LD is reversed. Hence, if selection changes direction without allowing linkage equilibrium to decay between epochs, the evolution of trait correlation takes longer. These episodes of counter-selection at the beginning of each epoch, owing to unresolved ‘wrong’ LD, are reflected in a characteristic saw-tooth pattern in figure 3b. If however the linkage equilibrium is established between opposing epochs of directional selection, the correlation between the traits evolves at about the same rate. This can be seen from the comparison of figures 2a and 3a, both showing positive co-selection of both traits. Figure 3a also shows slight ‘wrinkles’ in the curve. These are caused by the fact that at the beginning of each epoch, the LD necessary for selecting the rQTL allele has to be re-established, leading to a short period of slow selection. Figure 3c shows the same comparison but for the case of fluctuating selection in the corridor model, with equilibration between the epochs. The direction of the directional selection changes, while the stabilizing selection is constant. There is no effect of linkage relaxation in this model (not shown).

Figure 3. Evolution of genetic correlations under fluctuating selection. Four periods of co-selection were applied, each 100 generations long. In each epoch, the traits are simultaneously selected for higher mean values or lower mean values, respectively. All plots show three lines for simulations with different degrees of linkage ($L = 0.75$ (light grey), $0.8$ (black), $0.9$ (dark grey)).

(a) In this scenario, the linkage equilibrium between the rQTL and the rest of the genome is restored between each epoch of directional selection. This is equivalent to the assumption that between each epoch there is a period of stabilizing selection during which linkage equilibrium is restored (not simulated, but linkage equilibrium is enforced by averaging genotypic values). The rate of evolution of genetic correlations is similar to the scenario of sustained directional selection, only slightly delayed at the beginning of each epoch. The reasons for these delays are two. First, at the beginning of each epoch, LD has to be rebuilt that fuels the selection of the rQTL allele (see the plot of the rQTL allele frequency in the electronic supplementary material, figure S1a, which does not show a drop in frequency but only a slowdown in increase). Second, the relaxation of LD between epochs leads to a decrease of the population-level genetic correlation, which explains the instantaneous drop of correlation visible in the $L = 0.9$ curve at generation 100 and somewhat less so in the other curves and epochs. (b) Same scenario as in (a) except that the LD is not relaxed between epochs of directional selection. This scenario leads to a pronounced sawtooth pattern at the beginning of each new epoch. The reason is that initially, at the switch from one epoch to the next, the favoured rQTL allele is associated with genotypic values of the traits that cause lower fitness in the new selection regime. This ‘wrong’ LD is dissipated over a few generations and then selection for the favoured rQTL allele resumes. Overall, these effects do not significantly delay the evolution of the favoured form of genetic correlation. (c) Fluctuating selection in the corridor model, with equilibration between the epochs. The direction of the directional selection changes, while the stabilizing selection is constant. There is no effect of linkage relaxation in this model (not shown).
fluctuating directional selection on one trait and constant stabilizing selection on the second trait (i.e. corridor model). Other than the delays caused by the need to invert LD after each epoch of directional selection, this type of fluctuating selection is as effective as continued directional selection in shaping the correlation pattern of traits. Hence, rQTL selection under fluctuating environmental conditions requires some degree of recombination, because the LD needs to be inverted at the beginning of each epoch of directional selection. With full-linkage rQTL, selection is impossible under fluctuating environments (see the electronic supplementary material, figure S2). Note that the analytical solution in equation (2.2) and its consequences refer only to the persistent directional selection on both traits, but not to the corridor model or fluctuating selection.

4. DISCUSSION
Modelling the evolution of variational properties requires modelling multi-loci dynamics with gene interaction and multiple phenotypic characters under directional selection including the effect of LD. In most mathematical models most, if not all, of these complications are ignored in order to make the models tractable. The recent pioneering work by Jones et al. [30–32] focuses on stabilizing selection. In this project, we chose to make a number of alternative simplifications that preserve the presence of gene interaction and LD but use other abstractions. Specifically, we model the evolution of genes affecting two characters in the form of a phenotypic Lande equation and represent LD by an association between rQTL alleles and the genotypic values of the quantitative traits. To our knowledge, this model is the first to accommodate gene interaction, LD and directional selection and is still mathematically tractable. Certainly, the model is an abstraction that involves several simplifications. In particular, the G-matrix associated with a particular rQTL is kept constant and does not account for G-matrix dynamics caused by allele frequency changes. Furthermore, we assume that the rQTL alleles do not develop an additive effect. Epistatic rQTL can change their additive effect owing to background change [33]. We currently lack a general theory predicting the effect of epistasis on the multi-locus phenotype. Nevertheless, the parametrization of an analogous epistatic effect on univariate characters has been developed in the framework of the multi-linear model [33,34]. Within this framework, the overall effect of epistasis on the locus’ additive genetic effect on a single trait can be assessed. The parameter of the multi-linear model that measures the change in additive effect of the epistatic locus is the directionality of epistasis. Directionality describes how the additive effect of this locus on a phenotypic trait is modified on average owing to genetic interactions with the background loci. Negative values of directionality indicate diminished additive effect at the locus owing to epistasis, while the positive directionality indicates an increase of the additive effect owing to epistasis. A value close to zero implies that the effects of epistasis on the additive genetic effect of a locus cancel out in total across background loci. In this terminology, the assumption of the constancy of genetic effect across generations implies that the directionality of epistasis of the rQTL is zero for both phenotypic traits. A recent empirical study documented low net effect of epistasis on additive variance in mice intercross [35]. Additional studies of directionality are required to show how frequently this assumption holds.

Hence, at this stage we do not include all the complexities of the G-matrix dynamics. It is known that even with an additive model, the evolutionary dynamics of the G matrix is very difficult to model mathematically. The difficulties are further increased greatly if epistasis is taken into account. The benefits of doing so, however, could be quite limited. It is well understood that under sustained directional selection, the genetic variance approaches quasi-equilibrium between novel mutations and fixation of alleles [36]. This equilibrium is equal to the neutral mutation selection equilibrium, and thus is proportional to the mutational variance of the character. Furthermore, artificial directional selection has relatively small effects on genetic variance [37]. Another potential expansion of a current model is by including multiple rQTL, as in all documented cases, intertrait relationship is affected by more than a single rQTL. While our model manifestly is an abstraction, we suggest it is one that can be highly productive until a stronger mathematical approach is found that can deal with the complexities of multiple characters, polygenic inheritance with gene interaction and LD.

In conclusion, the model presented here shows that, given the presence of ‘neutral’ rQTL, directional selection can cause an adaptive change in the genetic covariance structure. The result is that the direction of highest genetic variance aligns with the direction of selection ([38]; see figure 1c). This mechanism requires the build-up of LD between the rQTL and the rest of the genome. Our simulations show that even strong recombination can sustain rQTL selection because recombination does not instantly destroy LD. Selection on rQTL can decrease variational constraints caused by trait covariance. If directional selection acts on only one character, and the other character is under stabilizing selection, lower covariance, i.e. the dissociation of traits, is favoured. Such increased variational autonomy of traits facilitates independent adaptation and differentiation of traits and can contribute to an increase in organismal complexity, i.e. the individualization of body parts. As mentioned in §1, this is what happened in the stem lineage of anthropoid primates, probably facilitating the limb shape diversity of higher primates [16]. The model furthermore also accommodates fluctuating directional selection, which seems more plausible than long-term directional selection. In fluctuating selection, recombination plays an important role in facilitating the evolution of trait correlations.

In this model natural selection acts directly on evolvability by selecting genotypes that increase the rate of response to directional selection. The idea that selection could directly act on evolvability is controversial [39], but in recent years computational models have shown that selection for evolvability is probably a generic property of all but the most simplistic genetic models ([40–45]; see also [46] for an empirical example). The present model, together with the demonstration of the existence of quasi-‘neutral’ rQTL variation [17,23,24], shows that selection may be able to modify the genetic architecture to increase evolvability and, specifically, increase the response rate to

Proc. R. Soc. B (2011)
directional selection. In essence, we show that through differential epistatic interactions, selection can produce the ‘right kind’ of heritable phenotypic variation to accomplish adaptive change from the pre-existing genetic, but phenotypically cryptic, variation.

Finally, what are the possible physiological mechanisms causing this kind of variation? Complete account of the possible mechanisms of variation in pleiotropy is beyond the scope of the current study, yet we include an example here. He & Zhang [47] found in yeast that most of pleiotropy is owing to a single gene product involved in different spatial or temporal contexts, each context resulting in a different phenotype. They furthermore found that the degree of pleiotropy correlates positively with the number of protein–protein interactions of the product [47]. It is plausible then that the variation in the relationship between the end-phenotypes of these pathways arises because genes interacting with the focal product in one pathway do not overlap with the genes involved in other pathways, leading to somewhat independent evolvability of the multiple functions at the pleiotropic locus. This was exemplified in a study of yeast by Kim et al. [22], in which the authors analysed the molecular basis of a pleiotropic mutation at CYST4 (cystathionine beta-synthase) that causes rust coloration of colonies growing on copper sulphate and differences in drug sensitivity [48,49]. The single nucleotide mutation was shown to lower the levels of cysteine, an essential amino acid involved both in a sulphur assimilation pathway and in a detoxication pathway. Low levels of cysteine cause increased assimilation of sulphur and therefore rust coloration, and sensitivity to multiple drugs. Different gene products interact with the focal product, cysteine, in different pathways. Some of the interacting genes involved are specific to individual pathways, hence mutations in these genes (i.e., different backgrounds) affect the two phenotypes separately, causing variation in the strength of the relationship between them, in spite of a common gene product involved in both. The rQTL found in this case are the epistatically interacting loci affecting only one of the functions (pathways) of the pleiotropic locus, and causing different levels of correlation in different genotypes. Selection on a particular intertrait relationship according to our model would therefore fix the genotype at the epistatic locus. More focus on the molecular mechanisms involved will reveal further possibilities promoting variation in pleiotropy.

We thank Bruce Walsh for the comments on the manuscript. M.P. acknowledges support from Erwin Schrödinger Postdoctoral Fellowship of Austrian Science Foundation (FWF). J.M.C. acknowledges funding from NSF (no. BCS-0725068) and NIH (nos AR053224, DK055736 and GM065509). Research in the GPW laboratory is funded by the John Templeton Foundation. The opinions expressed in this article are not those of the John Templeton Foundation.

APPENDIX A

We present here two conceptually distinct ways to model the effect of recombination on the mean trait value. The model we used in the study is a phenomenological model. We also explain a classical population-genetic alternative below.

In the phenomological model, we do not consider an explicit mechanism of how the genotypes of the next generation are created, but assume that the mean phenotype of the offspring genotypes of a mating of genotypes with different average phenotype is a simple mixture of the average parental phenotypes. For instance, let us assume that the average phenotype of the genotypes with $R_i$ is $\bar{z}_i$ and that of $R_j$ is $\bar{z}_j$. Now let us calculate the average phenotype of the $R_i$ genotype in the next generation owing to recombination.

In the offspring generation, $R_i$ genotypes can arise from two types of matings: (i) matings of type $R_i \times R_b$, and (ii) matings of $R_b \times R_i$. Under random mating, the first kind of pairing, $R_i \times R_b$ occurs with frequency $p^2$. In this case, all offspring will be of $R_i$ type and will have the average phenotype $\bar{z}_i$, where $p$ is the frequency of $R_i$ genotypes and $q = 1 - p$ is the frequency of the $R_b$ genotypes. In the second, matings of $R_b \times R_i$, case pairings will occur with a frequency $2pq$. Of the offspring with this pairing, one-half will be of $R_i$ type and these will have the average phenotype $L\bar{z}_i + (1 - L)\bar{z}_j$,

where $L$ is a phenomenological parameter of linkage. If $L = 1$, there would be no recombination and the $R_i$ genotypes would retain their average parental phenotype. Assuming $L = 0.5$ would mean that the $R_i$ genotypes, coming from a $R_b \times R_i$ mating, would have the same average phenotype as the $R_i$ genotypes coming from the same mating. This would mean that, within one generation, the LD between $R_i$ and $R_b$ genotypes would be completely relaxed. We know that LD does not relax within one generation and thus restrict our phenomenological linkage parameter to the range $[0.75, 1]$. If $L = 0.75$, the resulting average phenotype will be half way between the average phenotypes of the $R_i$ parents and the linkage equilibrium phenotype $(\bar{z}_i + \bar{z}_j)/2$. Hence, the assumption is that the highest level of recombination will decrease the LD by a factor of 0.5. Taking the average phenotype resulting from the $R_b \times R_i$ and the $R_i \times R_b$ matings together and normalizing it by the frequency of $R_i$ genotypes, $p$, we obtain $\bar{z}_i' = (p + qL)\bar{z}_i + q(1 - L)\bar{z}_j$,

and analogously for the average phenotype of the $R_j$ genotypes we obtain $\bar{z}_j' = (pL + q)\bar{z}_j + p(1 - L)\bar{z}_i$.

An alternative method of deriving an analogous result is to consider the mean effect loci as a block, with no recombination within, only recombination relative to the rQTL. This way the whole block becomes associated with an allele at the rQTL, and the model becomes a haploid two-locus model. This can be modelled in the conventional way, using the rate of recombination $c$, with parameter space $0 - 0.5$. Consider a haploid organism with two alternative gametes, $i$ and $j$. We treat mean-effect-loci as a composite locus $M$, and thus we consider two loci on each gamete: $R (= rQTL)$ and $M$. Given the rate $c$ of recombination between $R$ and $M$, the two-locus genotypes in the offspring generation have the following frequencies and phenotypes:

Proc. R. Soc. B (2011)
The genotype $i$ at the rQTL thus has the frequency $p^2 + pq = p$. The average phenotype of this genotype class is the weighted average of the phenotypes of two parental genotypes, which contribute to this offspring class (R,Mi, R,Mj).

$$
\bar{z}_i = \frac{1}{p} \left( p^2 \bar{z}_i + pq(1-c)\bar{z}_j + pq\bar{z}_j \right)
$$

$$
= p\bar{z}_i + q(1-c)\bar{z}_j + q\bar{z}_j,
$$

where $\bar{z}_j'$ stands for the similar mean offspring phenotype of the genotype class $i$ at the rQTL.

Interestingly, this method yields a similar equation to predict phenotypes after recombination, where the linkage parameter $L$ is replaced here by $(1-c)$, $c$ being the recombination rate. The only difference is that $1 - c$ has the parameter range $[0, 0.5, 1]$, while that of $L$ is $[0, 0.75, 1]$. This difference reflects the fact that the phenomological model implicitly allows mixing (recombination) among the mean effect loci.

### REFERENCES

1. Terentjev, P. V. 1931 Biometrische Untersuchungen ueber die morphologischen Merkmale von *Rana ridibunda* Pall. (Amphibia, Salientia). *Biometrika* 23, 23–51.
2. Berg, R. L. 1960 The ecological significance of correlation pleiades. *Evolution* 14, 171–180. (doi:10.2307/2405824)
3. Chetverikov, S. S., Barker, M. & Lerner, I. M. 1961 On certain aspects of the evolutionary process from the standpoint of modern genetics. *Proc. Am. Phil. Soc.* 105, 167–195.
4. Darwin, C. 1859 *On the origin of species*. London, UK: John Murray.
5. Olson, E. C. & Miller, R. L. 1958 *Morphological integration*. Chicago, IL: University of Chicago Press.
6. Armbruster, W. S., Di Stilio, V. S., Tuxill, J. D., Flores, T. C. & Velasquez Runk, J. L. 1999 Covariance and decoupling of floral and vegetative traits in nine neotropical plants: a re-evaluation of Berg’s correlation-pleiades concept. *Am. J. Bot.* 86, 39–55. (doi:10.2307/2656935)
7. Hansen, T. F., Armbruster, W. S., Carlson, M. L. & Pelabon, C. 2003 Evolvability and genetic constraint in *Dalsechamaria* blossoms: genetic correlations and conditional evolvability. *J. Exp. Zool. B Mol. Dev. Evol.* 296, 23–39.
8. Marroig, G., De Vivo, M. & Cheverud, J. M. 2004 Cranial evolution in salaks (*Pithecia, Platyrhinini*). II. Evolutionary processes and morphological integration. *J. Evol. Biol.* 17, 144–155. (doi:10.1046/j.1420-9101.2003.00653.x)
9. Hallgrímsson, B., Dorval, C. J., Zelditch, M. L. & German, R. Z. 2004 Craniofacial variability and morphological integration in mice susceptible to cleft lip and palate. *J. Anat.* 205, 501–517. (doi:10.1111/j.0021-8782.2004.0356.x)
10. Cheverud, J. 1995 Morphological integration in the saddle-back tamarin (*Saguinus fuscicollis*) cranium. *Am. Nat.* 145, 63–89. (doi:10.1086/285728)
11. Mezey, J. G., Cheverud, J. M. & Wagner, G. P. 2000 Is the genotype-phenotype map modular? A statistical approach using mouse quantitative trait loci data. *Genetics* 156, 305–311.
12. Leamy, L., Pomp, D., Eisen, E. J. & Cheverud, M. J. 2002 Pleiotropy of quantitative trait loci for organ weights and limb bone lengths in mice. *Physiol. Genet. 10*, 21–29.
13. Ehrich, T. H., Vaughn, T. T., Koreishi, S. F., Linsey, R. B., Pletscher, L. S. & Cheverud, J. M. 2003 Pleiotropic effects on mandibular morphology. I. Developmental morphological integration and differential dominance. *J. Exp. Zool.* 296, 58–79. (doi:10.1002/jex.b.9)
14. Albertson, R. C., Streelman, J. T., Koche, T. D. & Yelick, P. C. 2005 Integration and evolution of the chichlid mandible: the molecular basis of alternate feeding strategies. *Proc. Natl Acad. Sci. USA* 102, 16 287–16 292. (doi:10.1073/pnas.0506649102)
15. Young, N. M. & Hallgrímsson, B. 2005 Serial homology and the evolution of mammalian limb covariation structure. *Evolution* 59, 2691–2704.
16. Young, N. M., Wagner, G. P. & Hallgrímsson, B. 2010 Development and the evolvability of human limbs. *Proc. Natl Acad. Sci. USA* 107, 3400–3405. (doi:10.1073/pnas.0911856107)
17. Pavlicev, M., Kenney-Hunt, J. P., Norgard, E. A., Reman, G. C., Wolf, J. B. & Cheverud, M. J. 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.
18. Lenski, R. E. 1988 Experimental studies of pleiotropy and epistasis in *Escherichia coli*. I. Variation in competitive fitness among mutants resistant to virus T4. *Evolution* 42, 425–432. (doi:10.2307/2409028)
19. Lenski, R. E. 1988 Experimental studies of pleiotropy and epistasis in *Escherichia coli*. II. Compensation for maladaptive effects associated with resistance to virus T4. *Evolution* 42, 433–440. (doi:10.2307/2409029)
20. Templeton, A. R., Crease, T. J. & Shah, F. 1985 The molecular through ecological genetics of abnormal abdomen in *Drosophila mercatorum*. I. Basic genetics. *Genetics* 111, 805–818.
21. Templeton, A. R., Hollocher, H. & Johnston, J. S. 1993 The molecular through ecological genetics of abnormal abdomen in *Drosophila mercatorum*. V. Female phenotypic expression on natural genetic backgrounds and in natural environments. *Genetics* 134, 475–485.
22. Kim, H. S., Huh, J. & Fay, J. C. 2009 Dissecting the pleiotropic consequences of a quantitative trait nucleotide. *FEMS Yeast Res.* 9, 713–722. (doi:10.1111/j.1567-1364.2009.00516.x)
23 Cheverud, J. M., Ehrich, T. H., Vaughn, T. T., Koreishi, S. F., Linsey, R. B. & Pletscher, L. S. 2004 Pleiotropic effects on mandibular morphology. II. Differential epistasis and genetic variation in morphological integration. *J. Exp. Zool. B* **302**, 424–435.

24 Leamy, L. J., Pomp, D. & Lightfoot, J. T. 2009 Genetic variation in the pleiotropic association between physical activity and body weight in mice. *Genet. Select Evol.* **41**, 41–53. (doi:10.1186/1297-9686-41-41)

25 Lande, R. 1979 Quantitative genetic analysis of multivariate evolution, applied to brain: body size allometry. *Evolution* **33**, 402–416. (doi:10.2307/2407630)

26 Lande, R. & Arnold, S. J. 1983 The measurement of selection on correlated characters. *Evolution* **37**, 1210–1226. (doi:10.2307/2408842)

27 Wagner, G. P. 2010 The measurement theory of fitness. *Evolution* **64**, 1358–1376.

28 Wagner, G. P. 1988 The influence of variation and of developmental constraints on the rate of multivariate phenotypic evolution. *J. Evol. Biol.* **1**, 45–66. (doi:10.1046/j.1420-9101.1988.1010045.x)

29 Baatz, M. & Wagner, G. P. 1997 Adaptive inertia caused by hidden pleiotropic effects. *Theor. Popul. Biol.* **51**, 49–66. (doi:10.1006/tpbi.1997.1294)

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

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

32 Jones, A. G., Arnold, S. J. & Burger, R. 2007 The mutation matrix and the evolution of evolvability. *Evolution* **61**, 727–745. (doi:10.1111/j.1558-5646.2007.00071.x)

33 Hansen, T. F. & Wagner, G. P. 2001 Modeling genetic architecture: a multilinear theory of gene interaction. *Theor. Popul. Biol.* **59**, 61–86. (doi:10.1006/tpbi.2000.1508)

34 Carter, A. J. R., Hermisson, J. & Hansen, T. F. 2005 The role of epistatic gene interactions in the response to selection and the evolution of evolvability. *Theor. Popul. Biol.* **68**, 179–196. (doi:10.1016/j.tpb.2005.05.002)

35 Pavlicev, M., Le Rouzic, A. P., Wagner, G. P., Cheverud, J. M. & Hansen, T. F. 2010 Directionality of epistasis in a murine intercross population. *Genetics* **185**, 1489–1505. (doi:10.1534/genetics.110.118356)

36 Hill, W. G. 1982 Rates of change in quantitative traits from fixation of new mutations. *Proc. Natl Acad. Sci. USA* **79**, 142–145. (doi:10.1073/pnas.79.1.142)

37 Burger, R. 1993 Predictions of the dynamics of a polygenic character under directional selection. *J. Theor. Biol.* **162**, 487–513. (doi:10.1006/jtbi.1993.1101)

38 Schluter, D. 1996 Adaptive radiation along genetic lines of least resistance. *Evolution* **50**, 1766–1774. (doi:10.2307/2410734)

39 Lynch, M. 2007 The frailty of adaptive hypotheses for the origins of organismal complexity. *Proc. Natl Acad. Sci. USA* **104**(Suppl. 1), 8597–8604. (doi:10.1073/pnas.0702207104)

40 Bergman, A. & Siegal, M. L. 2003 Evolutionary capacitance as a general feature of complex gene networks. *Nature* **424**, 549–552. (doi:10.1038/nature01765)

41 Earl, D. J. & Deem, M. W. 2004 Evolvability is a selectable trait. *Proc. Natl Acad. Sci. USA* **101**, 11531–11536. (doi:10.1073/pnas.0404656101)

42 Kashtan, N. & Alon, U. 2005 Spontaneous evolution of modularity and network motifs. *Proc. Natl Acad. Sci. USA* **102**, 13773–13778. (doi:10.1073/pnas.0503610102)

43 Draghi, J. & Wagner, G. P. 2008 Evolution of evolvability in a developmental model. *Evolution* **62**, 301–315. (doi:10.1111/j.1558-5646.2007.00303.x)

44 Draghi, J. & Wagner, G. P. 2009 The evolutionary dynamics of evolvability in a gene network model. *J. Evol. Biol.* **22**, 599–611. (doi:10.1111/j.1420-9101.2008.01663.x)

45 Parter, M., Kashtan, N. & Alon, U. 2008 Facilitated variation: how evolution learns from past environments to generalize to new environments. *PLoS Comput. Biol.* **4**, e1000206. (doi:10.1371/journal.pcbi.1000206)

46 Allen, C. E., Beldade, P. & Brakefield, P. M. 2008 Differences in the selection response of serially repeated color pattern characters: standing variation, development, and evolution. *BMC Evol. Biol.* **8**, 94. (doi:10.1186/1471-2148-8-94)

47 He, X. & Zhang, J. 2006 Toward a molecular understanding of pleiotropy. *Genetics* **173**, 1885–1891. (doi:10.1534/genetics.106.060269)

48 Fay, J. C., McCullough, H. L., Sniegowski, P. D. & Eisen, M. B. 2004 Population genetic variation in gene expression is associated with phenotypic variation in *Saccharomyces cerevisiae*. *Genome Biol.* **5**, R26. (doi:10.1186/gb-2004-5-4-r26)

49 Kim, H. S. & Fay, J. C. 2007 Genetic variation in the cysteine biosynthesis pathway causes sensitivity to pharmacological compounds. *Proc. Natl Acad. Sci. USA* **104**, 19 387–19 391. (doi:10.1073/pnas.0708194104)