Research

THE ROYAL SOCIETY
PUBLISHING

Evolutionary trade-offs and the structure of polymorphisms

Hila Sheftel[1], Pablo Szekely[1], Avi Mayo[1], Guy Sella[2] and Uri Alon[1]

[1]Department Molecular Cell Biology, Weizmann Institute of Science, Rehovot 76100, Israel
[2]Department of Biological Sciences, Columbia University, New York, NY 10027, USA

(iD) UA, 0000-0001-6903-9956

Populations of organisms show genetic differences called polymorphisms. Understanding the effects of polymorphisms is important for biology and medicine. Here, we ask which polymorphisms occur at high frequency when organisms evolve under trade-offs between multiple tasks. Multiple tasks present a problem, because it is not possible to be optimal at all tasks simultaneously and hence compromises are necessary. Recent work indicates that trade-offs lead to a simple geometry of phenotypes in the space of traits: phenotypes fall on the Pareto front, which is shaped as a polytope: a line, triangle, tetrahedron etc. The vertices of these polytopes are the optimal phenotypes for a single task. Up to now, work on this Pareto approach has not considered its genetic underpinnings. Here, we address this by asking how the polymorphism structure of a population is affected by evolution under trade-offs. We simulate a multi-task selection scenario, in which the population evolves to the Pareto front: the line segment between two archetypes or the triangle between three archetypes. We find that polymorphisms that become prevalent in the population have pleiotropic phenotypic effects that align with the Pareto front. Similarly, epistatic effects between prevalent polymorphisms are parallel to the front. Alignment with the front occurs also for asexual mating. Alignment is reduced when drift or linkage is strong, and is replaced by a more complex structure in which many perpendicular allele effects cancel out. Aligned polymorphism structure allows mating to produce offspring that stand a good chance of being optimal multi-taskers in at least one of the locales available to the species.

This article is part of the theme issue 'Self-organization in cell biology'.

1. Introduction

Organisms show prevalent genetic differences called polymorphisms. Polymorphisms, together with environmental and epigenetic effects, cause much of the variation in traits between individuals. Each polymorphism typically affects multiple traits (pleiotropy), and each trait is usually affected by many different polymorphisms [1,2].

Understanding the origin and distributions of polymorphisms is important in evolutionary and quantitative genetics. Most theoretical approaches to the evolution of polymorphisms employ the classic picture of the fitness landscape: a genotype determines phenotype, and the phenotype determines fitness [3] (figure 1*a*). Natural selection leads to phenotypes that maximize fitness, provided that there is sufficient time, genetic variation, and that populations and selection pressures are large enough to overcome genetic drift [5]. In this approach, the shape of the fitness landscape, in particular the slopes near the maximum, influences the distribution of phenotypes and genotypes in the population. For example, simulations show that a fitness ridge can lead to mutation-selection balance in which traits vary along the ridge [6,7].

Here, we consider an extension of the fitness landscape picture for cases in which fitness derives from an organism's performance at *multiple tasks*. The need to perform multiple tasks introduces an inherent trade-off, because in most cases no single phenotype can be optimal at all tasks. For example, a bird
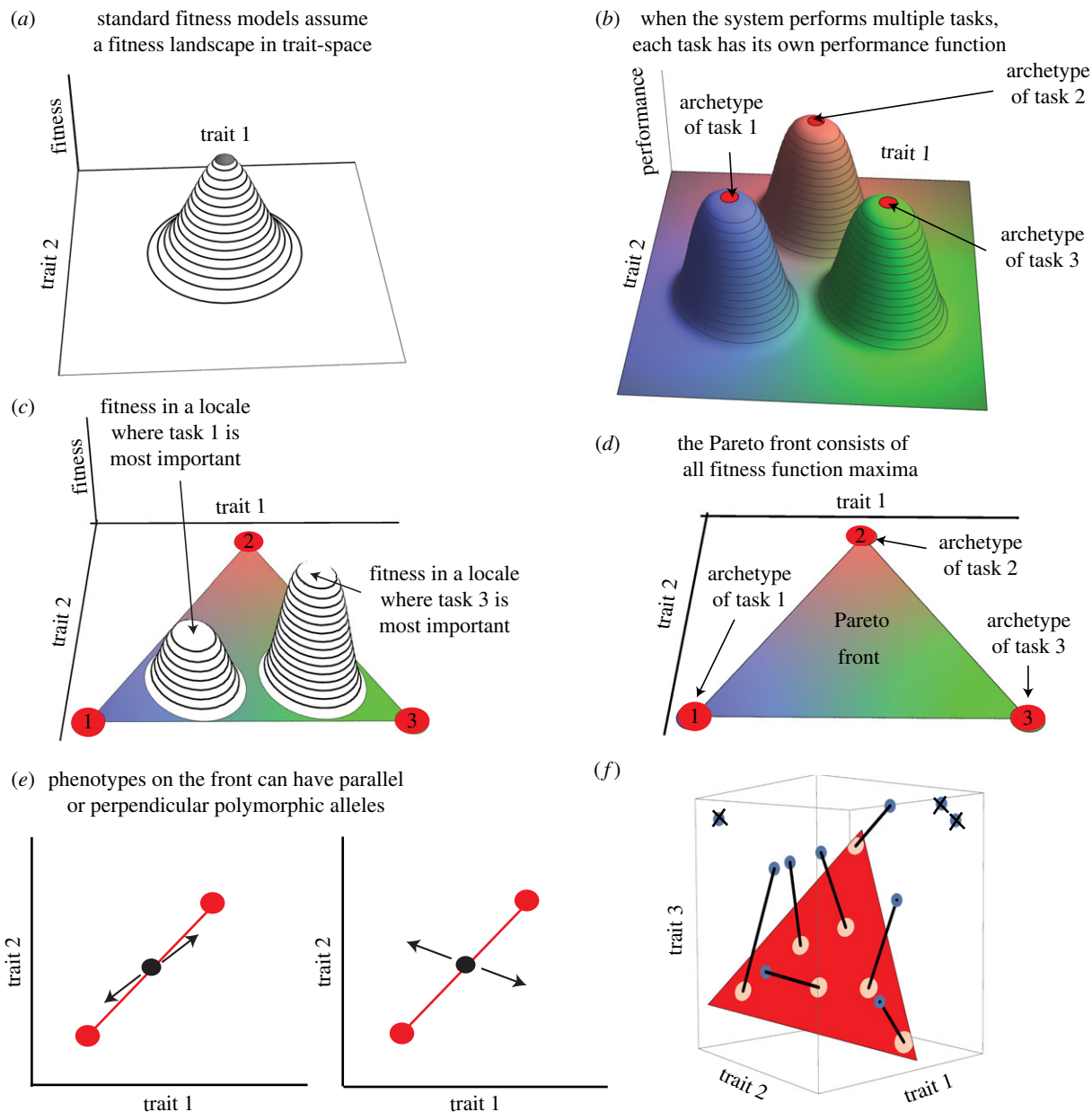
Figure 1. Trade-off between tasks leads to phenotypes arranged along polygons in trait space (*a*) Standard approaches assume a fitness landscape in the space of phenotypic traits. Selection tends to favour phenotypes near the maximum. (*b*) When the system needs to perform several tasks, each task has a performance function in trait space, whose maximum is called the archetype. Each coloured hill represent a performance function $P_i(T)$ for task i, where T is the vector of traits, and fitness at locale q is an increasing function of all performance functions $F_q(P_1(T), \ldots, P_m(T))$ (*c*) In an environment where task 1 is more important, fitness is maximized at a phenotype close to archetype 1; In an environment where task 3 is more important, fitness is maximized at a phenotype closer to archetype 3. Fitness maxima in the two environments are different. (*d*) Fitness maxima in all possible environments in which fitness is an increasing function of performance in the three tasks lie in the full triangle whose vertices are the three archetypes—known as the Pareto front—when the conditions of Shoval *et al.* [4] are fulfiled. (*e*) Possible polymorphism structures that keep a phenotype on the front: either mutation effects are parallel to the front (left), or mutations whose components perpendicular to the front cancel out (right). (*f*) Schematic of the selection process described in the Results section, for the case of three tasks that define a triangle whose vertices are the three archetypes. Locales and individuals best at the locales are sequentially removed into a survivor list, and remaining individuals are removed. Survivors make up the next generation. Here, we consider 7 locales ($N = 7$); phenotypes that maximize fitness in these locales are connected to the relevant locale with a black line.

may eat both seeds and pollen. However, cracking seeds requires a beak shaped like a plier whereas pollen requires a pincer-like beak [8].

In such cases, the 'genotype → phenotype → fitness' scheme should be modified to include performance at different tasks [9]: genotype → phenotype → performance at task 1,2, . . . → fitness). The genotype determines phenotype, which is defined by a set of traits T such as morphological parameters (e.g. beak width and length), gene expression levels or enzyme activities. The phenotype, in turn, determines the performance at different tasks: The performance at task i is given

by the performance function $P_i(T)$ (figure 1*b*). Fitness is an increasing function of the performances, $F(P_1(T), P_2(T) \ldots)$. The fitness, F, in a given locale is given by a (potentially non-linear) weighting of the performances in the different tasks (figure 1*c*). Therefore, each locale has its own fitness function that combines the performances in the tasks.

This picture of trade-offs between tasks has been shown to result in a simple geometry of phenotypes in the space of traits (trait space) [4]. Under quite general assumptions, the optimal phenotypes in all possible environments fall within polytopes in trait-space (polytopes are the generalization of polygons to

any dimension, such as lines, triangles, tetrahedra, etc.). The vertices of the polytopes are the phenotypes that are optimal for a single task. These vertices are called *archetypes*, and the polytopes they define are the *Pareto fronts*.

Each point on the Pareto front corresponds to a phenotype that maximizes fitness in a locale that requires a specific degree of specialization in each of the tasks (figure 1c). For example, a trade-off between two tasks leads to phenotypes along a line segment, with the two archetypes at either end; generalists lie in the middle. A trade-off between three tasks leads to a full triangle (figure 1d), four tasks lead to a tetrahedron and so on. Such straight edged polytopes occur in coordinate systems in which performance declines with a distance metric [10]. If trait axes are nonlinearly transformed (e.g. trait → trait^2), the polygons have curved edges, but still preserve the distinct vertices at the archetypes optimal for each task.

Evidence for such lines, triangles and tetrahedra was found in morphological datasets [4,11,12]; animal life-history traits [13]; locomotive behaviour [14,15] and gene expression data [4,16–18]. The polytopes offer a way to deduce the tasks from the data, by observing the special features of organisms near each archetype. In the studies cited above, the archetypes found by fitting the data to polytopes [16] revealed clues about the tasks at play in each case. For example, primary tasks for *E. coli* gene expression are growth and survival [4]. Because this theoretical approach allows inference of tasks from a multi-trait dataset, it is called Pareto task inference (ParTI).

Up to now, work on the ParTI approach has not considered its genetic underpinnings. Here, we address this by asking how the polymorphism structure of a population is affected by evolution under trade-offs between multiple tasks. We simulate a multi-task selection scenario, in which the population evolves to the Pareto front: the line segment between the two archetypes or the triangle between three archetypes. We find that the polymorphisms that become prevalent in the population have phenotypic effects that align with the Pareto front. Epistatic effects between prevalent polymorphisms are also parallel to the front. This polymorphism structure allows rapid evolution to new environments that require the same tasks at different weightings. It also provides a mechanism that allows mating to produce offspring that stand a good chance of being optimal multi-taskers. Alignment to the front is reduced when genetic drift is strong, and is replaced by a more complex structure in which perpendicular effects of alleles collectively cancel out.

2. Material and methods

We assume that performance at each task decays with distance r to the archetype [4,19]. Performance functions were $P_i(T) = P_i(r(T, A_i))$, where distance r from the archetype A_i is an inner-product norm $r(T, A_i) = (T − A_i)^T Q_i (T − A_i)$ with a positive-definite matrix Q_i (Euclidean distance is obtained for $Q_i = I$). Unless otherwise noted, all simulations used $Q_i = I$. We chose population sizes N for which simulation duration for 10N generations was feasible, namely up to $N = 2000$.

3. Results

(a) A selection scheme for multi-task evolution

We consider a setting where fitness is determined by the performance at L different tasks. Phenotypes are described as vectors T in a trait space, where each axis corresponds to a quantitative trait. The performance functions for the L tasks are $P_i(T)$ for $i = 1 \ldots L$. The maxima of these performance functions are the archetypes, $A_1 \ldots A_L$, the phenotypes optimal at each task, and performance at each task decays with distance from the archetype [4,19] (Methods).

To address evolution under a trade-off between the tasks, we developed a selection process as follows: we consider N discrete locales, such as territories or nesting sites, each of which can be occupied by a single individual (results apply also when locales can be occupied by several individuals, electronic supplementary material, §1). The locales differ in environmental factors, and therefore the different tasks are more or less important for overall fitness at each locale. This heterogeneity can result from geographic clines, patchy environments, variation in other species, and so on. Hence, each locale has its own fitness function (individual selection surface) with locale-specific weights for the different performance functions: $F_{\{w_i\}} = \Sigma_i w_i P_i$, where the weights w_i are positive (similar results are found for fitness functions that are non-linear in the performances, electronic supplementary material, §2). In the simulations shown below, we sample w_i uniformly with $\Sigma w_i = 1$. Other distributions for w_i yield the same qualitative results (electronic supplementary material, §3).

The weights w_i correspond to the fitness contribution of each task in the locale. With this mathematical description, the phenotype that maximizes fitness in each locale can be shown to fall within trait space in the polytope whose vertices are the archetypes [4], namely the Pareto front: a line for two tasks, a triangle for three tasks and so on.

Each round of the simulation begins with N diploid individuals, each in a different locale. We randomly choose kN pairs ($k > 1$) and mate each pair with recombination to generate one offspring per pair (simulations where mating probability depends on fitness yield the same qualitative result, electronic supplementary material, §4). Offspring gain new mutations with probability μ.

The kN offspring compete for the locales as follows. We choose a locale at random, select the individual with highest fitness according to the fitness function for that locale. We put that locale and individual aside in a 'survivor list', and repeat the process with remaining individuals and locales until all sites are filled. The remaining individuals are removed, and the N survivors make up the next generation (figure 1f). Selecting individuals stochastically according to their fitness yields the same quantitative results, electronic supplementary material, §5. As the parameter k increases, the competition for the sites intensifies and drift effects become smaller.

This selection scheme acts on individuals represented by a genome composed of a set of mutant alleles. Each mutation moves the phenotype in trait-space. Mutation effect is a randomly oriented vector whose length is drawn from an exponential distribution with mean $q = 1$, which is 1% of the distance in trait space between the archetypes. Mutation effects are additive (non-additive epistasis is addressed below). The number of new mutations introduced in an individual per generation is Poisson distributed with mean μ. For recombination, we assume an infinite site model [20,21] and free recombination [22] (for more details, see electronic supplementary material, §6).

We also simulated variants of this model as follows: (i) linkage, in which recombination was done by swapping chromosomes at a single recombination spot (ii) No
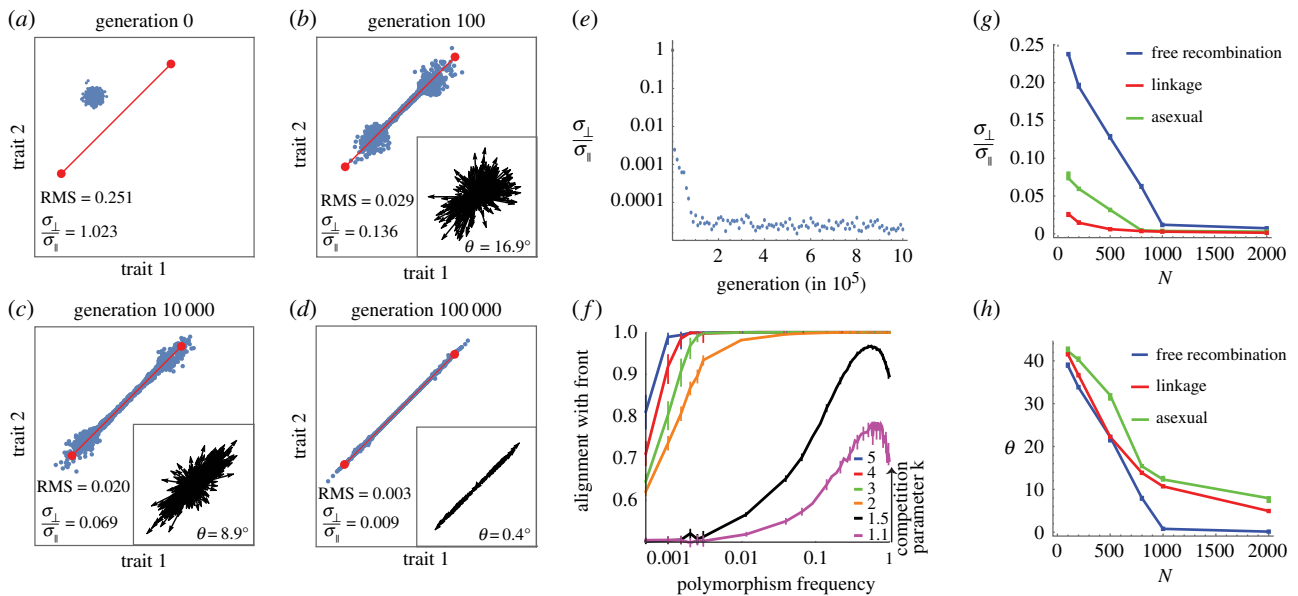
Figure 2. Phenotypes converge to the Pareto front by means of polymorphisms whose effects align with the front. (a–d) Snapshot of the phenotypes at different generations in a simulation, for the case of two tasks. The Pareto front is shown in red. Insets: The phenotypic effects of all polymorphisms present in at least 1% of genomes increasingly align with the Pareto front. Each arrow represents the phenotypic effect of one polymorphism (magnified for illustration purposes). At generation zero, no mutation is present at greater than 1% of the genomes. Simulation parameters: $N = 1000$, $\mu = 0.05$, $k = 2$, free recombination. Median angle θ of polymorphisms relative to the front was $17°$, $9°$, $0.4°$, for b–d, respectively. The ratio between perpendicular and parallel standard deviation of phenotypes is mentioned in each panel a–d, along with the phenotypic RMS distance from the front, normalized by the distance between archetypes. Axes are traits 1 and 2. (e) The ratio of the perpendicular and parallel standard deviation of phenotypes with respect to the front is 1 for the initial population at generation 0, decreases with generations, and begins to plateau after approximately 100 000 generations. (f) Polymorphisms that are aligned with the front tend to increase in frequency (log-linear scale) and with competition k. Simulation parameters are as above except k that varies as indicated. Error bars represent 95% confidence intervals from bootstrapping. Alignment is defined as $A = \langle|a_{\parallel}|\rangle^2/(\langle|a_{\parallel}|\rangle^2 + \langle|a_{\perp}|\rangle^2)$ where $\langle|a_{\parallel}|\rangle$ and $\langle|a_{\perp}|\rangle$ are the mean parallel and perpendicular component of the mutation effect vectors in each bin of mutation frequency. $A = 1$ and 0.5 occur when mutations are completely aligned or randomly oriented, respectively. Note that data are for frequencies less than 1, at frequency $= 1$ the mutation becomes fixed in the population. (g,h) Phenotypic standard deviation ratio $\sigma_{\perp}/\sigma_{\parallel}$ (g) and median angle θ (h) to the front, as a function of population size N, for three different recombination schemes: free recombination, linkage and asexual mating. In all simulations presented, $\mu N = 50$, and competition parameter $k = 2$. Error bars represent 95% confidence intervals from bootstrapping.

mating: asexual populations with a single chromosome and no recombination.

(b) Phenotypes in the population become spread along the Pareto front

We begin with simulations with two tasks, which define a Pareto front in the shape of a line segment between the two archetypes. The initial condition is a cloud of phenotypes that are off the Pareto front (figure 2a). For all parameters simulated here ($N\mu = 10$–50, $N = 100$–2000, $k = 1.1$–5), we find that within a few tens of generations the population spreads along the Pareto front. The distribution around the line becomes narrower as generations pass (figure 2b–e).

Regardless of parameters, phenotypic standard deviation parallel to the front $\sigma_{\parallel}$ is much larger than perpendicular standard deviation $\sigma_{\perp}$. For example, for $N = 1000$ and $k = 2$, we find after $100\,N$ generations $\sigma_{\perp}/\sigma_{\parallel} = 0.012 \pm 0.0003$. Similar results are found also when varying the shapes of the performance function contours ($Q_i \neq I$), and for linkage and asexual mating, as described in figure 2g and electronic supplementary material, §7.

(c) Prevalent polymorphisms have effects parallel to the Pareto front

We next asked about the polymorphism structure of the evolved population. In principle, there are two possible

structures that give rise to phenotypes along the front: either mutation effects are parallel to the front, or there exists pairs of mutations or higher-order combinations whose components perpendicular to the front cancel out (figure 1e).

We measured alignment with the front using the angle θ of each allele effect with the front. We find that while newly arising mutations are isotropic (median angle $= 45°$), polymorphisms that persist in the population (greater than 1% of the population) are closely aligned with the Pareto front (figure 2b–d insets). For example, the median angle of these polymorphisms is $0.80° \pm 0.03°$ after $100\,000$ generations, for typical simulation parameters ($N = 1000$, $k = 2$, $\mu = 0.05$). The more aligned a polymorphism is with the front, the more frequent it is in the population (figure 2f, electronic supplementary material, S8 and S9).

We next asked how genetic drift affects the alignment. We find that the smaller the population (the larger the drift), the weaker the alignment (figure 2h, electronic supplementary material, §8). Similarly, alignment is weaker when the competition parameter k is small, again a situation with larger drift (figure 2f). Drift effects seem to become important at population sizes smaller than a few hundreds.

We also asked about the effects of recombination on alignment. Alignment is reduced at a given population size when free recombination is replaced with single-site recombination (linkage) or in the case of asexual mating (no recombination) (figure 2h). At small populations sizes alignment is poor (greater than $20°$ at $N = 500$, for both asexual mating and
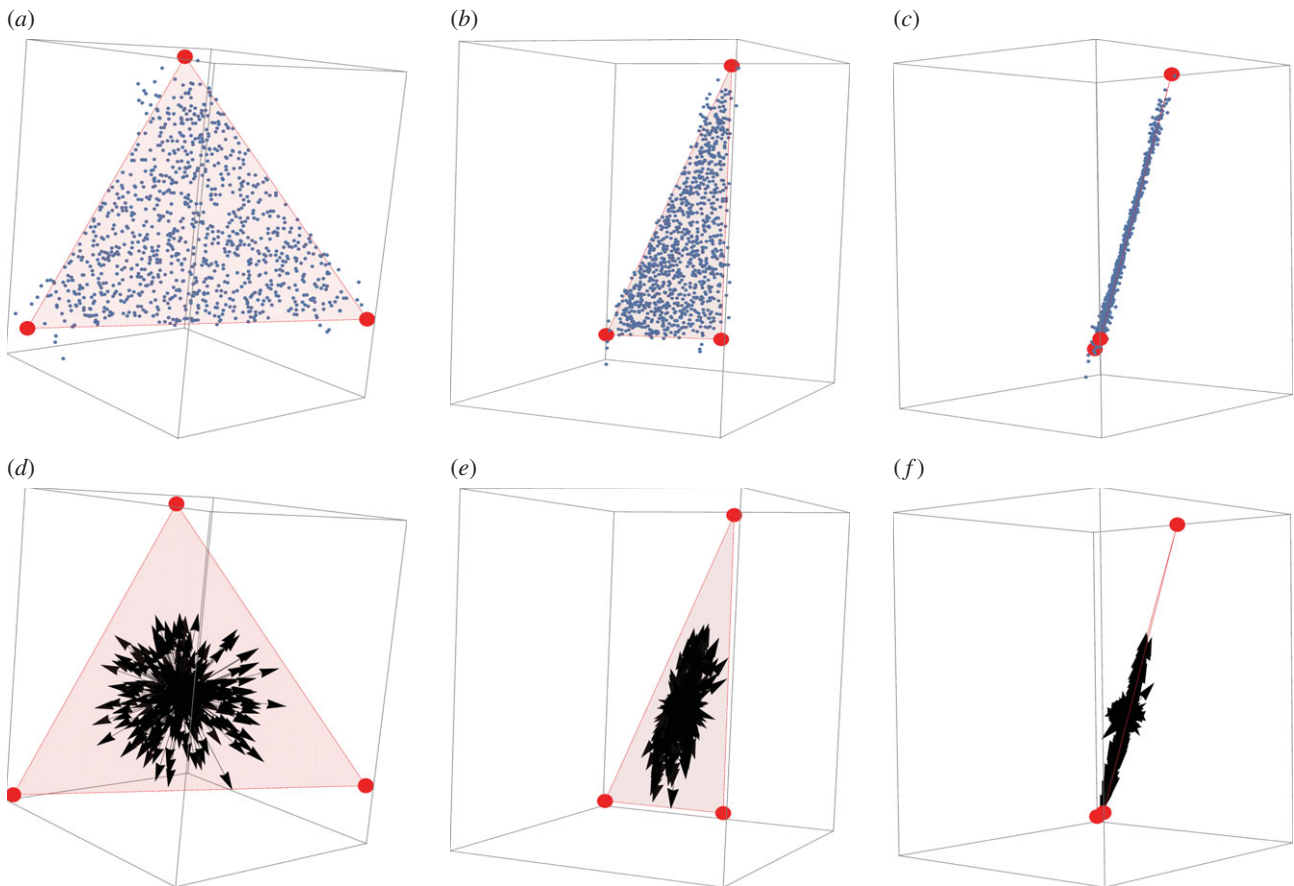
Figure 3. For three tasks, polymorphisms align with the plane of the triangular Pareto front (a–c) Phenotypes (blue dots) lie close to the Pareto (red triangle) front after 100 000 generations. Panels show the data from three different views. Simulation parameters are $N = 1000$, $\mu = 0.05$, $k = 5$, free-recombination. (d–f) Effects of polymorphisms present in at least 1% of the genomes (black arrows) align with the Pareto front (effects magnified by 10-fold for visualization). Each arrow represents the phenotypic effect of one polymorphism. Panels show the data from three different views.

linkage) but the phenotypes are still are close to the front ($\sigma_\perp / \sigma_\parallel < 0.05$), indicating that the polymorphisms have a complex structure in which perpendicular effects cancel out. Nevertheless, simulations with linkage or asexual mating still showed good alignment (angle $< 10°$) at large population sizes.

We conclude that regardless of linkage or asexual mating, at large enough population sizes common polymorphisms align with the front in these simulations, whereas at small population sizes alignment is weaker and instead polymorphism structure is more complex with cancelling perpendicular effects.

## (d) Phenotypic effects of polymorphisms align with moderately curved Pareto fronts

We also tested curved Pareto fronts which occur when the performance functions have eccentric contours ($Q_i \neq I$) that point at an angle with respect to each other [10], as shown in electronic supplementary material, §10 and figure S13. When the front curvature is mild, polymorphisms are still aligned with the local front (for example, when $k = 2$ and curvature is as shown in electronic supplementary material figure S11C, the median polymorphism angle is $6.1° \pm 0.1°$); this alignment is reduced at high front curvatures and high competition (for example, when $k = 5$ and curvature is higher as shown in electronic supplementary material figure S12B, the median angle is $24.1° \pm 0.2°$) (electronic supplementary material figure S11–13 and §10).

(e) Polymorphism effects align with a triangular Pareto front in the case of three tasks

We asked whether these conclusions apply also to higher numbers of tasks. We simulated evolution under three tasks, which results in a Pareto front shaped as a triangle whose vertices are the three archetypes. We used a three-dimensional trait space, with the triangle oriented at angle with respect to all trait axes. We find that, again, populations rapidly converge to the triangle (figure 3a–c). Prevalent polymorphisms align with the plane of the triangle (median angle approximately $2°$, figure 3d–f). This effect is seen also when mating probability depends on fitness (electronic supplementary material, §11).

(f) Epistatic effects between the two copies of the same allele align with the Pareto front

We next consider the effects of trade-offs on epistasis. We begin with the non-additive interaction between two alleles of the same gene [23–25], which is related to the phenomenon of dominance. We modelled this epistasis by assigning to each mutation i a randomly oriented vector d_i, such that the heterozygote mutant effect is m_i and the homozygote is $2m_i + d_i$. We find that as before, the main effects of prevalent polymorphisms m_i align with the front. Importantly, the epistatic effects d_i of common polymorphisms also align with the front (electronic supplementary material, §12). Those with epistatic effects off of the front are selected
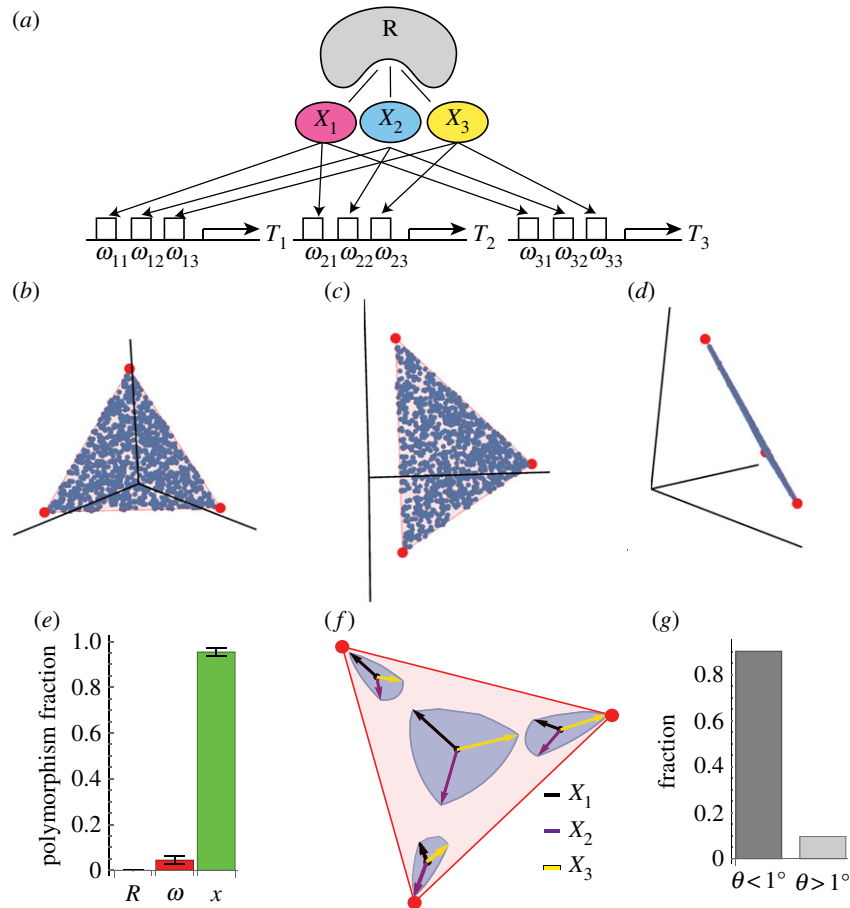
Figure 4. Epistasis due to a molecular mechanism evolves to keep phenotypes within a triangular front (*a*) A mechanism based on bacterial transcription considers genes whose promoters are regulated by 3 regulators (sigma factors) X_i that compete over a limiting factor R (RNA polymerase). The effect of X_i on the promoter of gene j is ω_{ij}. (*b–d*) Phenotypes (blue dots) reside on the Pareto (red triangle) front after 100 000 generations with this mechanism. Panels show the data from three different views. Each mutation changed one of the parameters X_i, ω_{ij} or R by multiplying the parameter by number drawn from a lognormal distribution LN(1,1). (*e*) Fraction of prevalent polymorphisms (greater than 1% of genomes) in each of the parameters R, ω_{ij} (summed over all ωs) and X_i (summed over X_1, X_2, X_3). Median and median deviation over 100 different simulations is shown at 100 000 generations. (*f*) Effect of mutations that change X_i by 10%. Arrows are 10% changes in X_1 (black arrow), X_2 (purple arrow) or X_3 (yellow arrow), and the blue contour is the boundary of the effects of simultaneously changing X_1, X_2 and X_3 so that the total change is 10%. Mutation effects are magnified by 12 for illustration. Mutations show epistasis, by having different effects at different genetic backgrounds that correspond to four different phenotypes on the front. (*g*) Fraction of angles with respect to the plane of the triangle that are smaller (left) or bigger (right) than 1°, for all polymorphisms present at greater than 1% of the genomes. Each polymorphism was evaluated in all of the genomes in which it appears, because the angle can vary due to epistasis. Results are based on simulation with parameters $N = 1000$, $\mu = 0.05$, $k = 5$, free-recombination. Each mutation changed one of the parameters X_i.

against. Thus epistasis tends to align with the front, provided that drift is not too large.

(g) Epistasis in a molecular mechanism that generates triangular fronts

In addition to epistasis between the alleles of the same gene, we studied epistasis between different genes. In this case, the effects of each polymorphism depend on the genetic background. To study this, we focus on epistasis due to non-linear interactions within a molecular mechanism inspired by gene expression [4]. This molecular mechanism is of interest because it occurs in bacteria, and naturally provides scope for a triangular Pareto front.

In the model, genes are regulated by three regulators X_i that compete over a limiting factor R. For example, in bacteria, three sigma factors compete over RNA polymerase [26,27]) (figure 4*a*). The traits are the expression of different genes. The genes are regulated by the three sigma factors

bound to RNA polymerase which bind to sites in the promoter of the gene. The effect of regulator i on gene j expression is ω_{ij}. Thus, expression of gene j is $T_j = R\Sigma_{i=1}^{3}\omega_{ij}(X_i/\Sigma_m X_m)$ (electronic supplementary material, §13). The competition between X_i for binding to R leads to the nonlinear term $X_i/\Sigma_m X_m$ which causes epistasis between allelic variants.

We simulated evolution in this model (in bacteria such as *E. coli*, the three tasks can be growth, survival and motility regulated by the σ-factors σ_{70}, σ_S and σ_A [26,27]). Each mutation varies one of the biochemical parameters: X_i, ω_{ij} or R. Phenotypes evolve to the triangular front (figure 4*b–d*).

Importantly, we find that almost all of the prevalent polymorphisms affect the levels of the regulators X_i; there are almost no prevalent polymorphisms that affect the levels of the other biochemical parameters, the promoter strengths ω_{ij} or RNA polymerase levels R (figure 4*e*), once the mutations that set ω_{ij} become fixed (at around 1000 generations, electronic supplementary material, §14).

Thus, evolution first encodes the coordinates of the three archetypes in the promoter sequences that set the weights ω_{ij}. The position on the triangle is then given by the relative values of the regulators X_i. For example, phenotypes at the vertices of the triangle have one regulator X_i high and the rest very low.

The polymorphisms in X_i have phenotypic effects that depend on the genetic background (figure 4f) due to the epistasis. For genetic backgrounds near the centre of the front, polymorphisms in X_i move the phenotype in all directions along the plane of the triangle; near the vertices, they move the phenotype in a tapered way that fits into the corner of the triangle, preventing phenotypes from leaving the triangle. None of these polymorphisms has a sizable component that moves off of the plane of the triangle (figure 4g). By contrast, the mutations that are actively selected against and hence do not reach high prevalence—those in the parameters ω_{ij} (promoter mutations) and R (RNA polymerase)—have effects that move the phenotype off of the plane the triangle. For example, increasing a weight ω_{ij} moves expression of gene i up, without affecting the other genes, a move which is off of the plane of the triangle. Altering the level of the limiting substrate R changes expression of all genes at once, a direction that is also off of the plane of the triangle. Expression levels of R and of genes will presumably be determined by the benefit and cost of production and maintenance. Such cost is included in the performance functions—increasing expression of genes carries a cost that, at high enough expression levels, reduces performance [28,29]. This demonstrates how a regulatory mechanism focuses the selection of mutations to certain components (regulator activity) and not to others (promoters, RNA polymerase). The common polymorphisms have epistatic interactions that keep them within a sharp triangle of phenotypes [30]. The same idea can be generalized to any number of tasks by using more regulators.

(h) Polymorphism structure under trade-offs allows rapid selection along the Pareto front

The polymorphism structure found above has implications for artificial selection towards desired traits. To explore this, we exposed populations that have evolved for 100 K generations to selection towards a given phenotypic target (electronic supplementary material, §15). We compared selection to a target aligned with the original front to selection to a target that is not aligned with the front. Each generation, we selected the fraction p of the population closest to the target.

We find that evolution to a target aligned with the original front is much faster and results in phenotypes that move farther from the original phenotypes than selection perpendicular to the original front (figure 5a,b and electronic supplementary material, §15 and figure S17B). For example, for typical 3-task simulation parameters ($N = 1000$, $k = 5$, $\mu = 0.05$), and artificial selection parameter p in the range 0.1–0.9, response to selection aligned with the original front is 9-fold to 17-fold larger than the response to selection perpendicular to it.

Evolution to a target off the front results in rapid motion to the point on the front that is closest to the target (the projection of the target on the plane of the triangle). This is followed by much slower evolution off of the front that stops when perpendicular genetic variation is depleted (figure 5a).

Simulations of selection in the case of two tasks (figure 5b) are qualitatively similar to experiments on *Bicyclus anynana* butterfly eyespot size [31], in which artificial selection parallel to the observed suite of variation leads to stronger evolutionary change than selection perpendicular to the suite of variation (figure 5c).

Additional simulations show that prolonged selection perpendicular to the original front can generate a population that is spread along a hyperplane parallel to the original front (plane for 3 tasks and a line for 2 tasks) (figure 5d and electronic supplementary material, §15 and figure S17d). A similar effect has been observed in experiments on radish morphology (figure 5e) [32,34].

(i) The present predictions are supported by a range of experimental data

To further test the present conclusions requires experiments that perturb molecular pathways and measure the resulting phenotypic changes, in relation to the natural suite of variation. A set of studies that exemplify this approach considered the pathways that shape the beak of Darwin's finches [33,35,36].

The normalized beak dimensions of different Darwin's ground finch species fall approximately on a plane (first two principle components account for 99.6% of the variation, figure 5f), and within that plane on a triangle corresponding to the tasks of eating large seeds, small seeds and pollen/nectar [4,37]. Perturbations of the major beak morphogenic pathways, BMP, calmodulin and premaxillary-bone modification factors (TGFβIIr, β-catenin and Dkk3), have phenotypic effects that are pleiotropic in a way that is approximately aligned with this plane [33] (figure 5f). We hypothesize that polymorphisms that affect the expression of these factors are likely to be prevalent and underlie the variation of the beak morphology along the Pareto front.

One can extend this framework to consider different species in a taxon that have comparable traits and share the same tasks. In this case, a prediction of additive, aligned polymorphism structure is that populations within a species will have variation that is aligned with the front defined by the variation of different species [4,38]. A recent example on bird toe-bone proportions showed that different bird species fall on a triangle. The three tasks in this case are grasping, walking and scratching/rapturing (figure 5g) [11]. Populations of chicken and zebra finch individuals fall in a flattened cloud aligned with the front defined by the other bird species, as predicted (figure 5g).

4. Discussion

We find that polymorphisms that persist under multi-task selection have phenotypic effects oriented along the Pareto front rather than perpendicular to the front. Similarly, polymorphisms that persist have epistatic effects that tend to keep phenotypes on the front.

We also tested the effects of genetic drift on the polymorphism structure. Small population sizes and low competition, both situations with large drift, lead to weaker alignment with the front.

Polymorphism structure is also affected by linkage, which can cause alleles to be carried together. We therefore studied the effect of linkage and asexual mating on the polymorphism
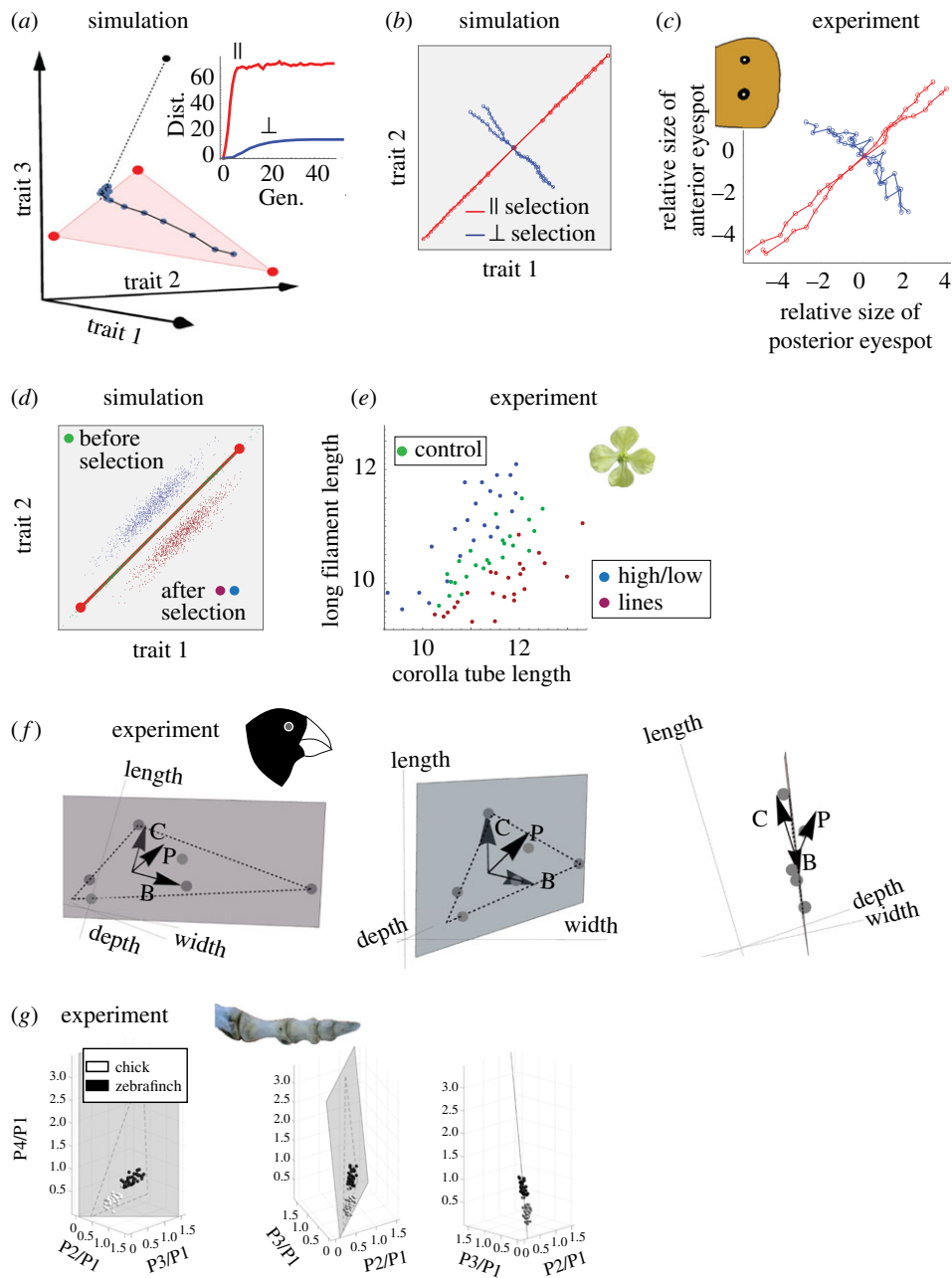
Figure 5. Simulations on artificial selection and population variation agree with a range of experimental results (*a*) Response to selection towards a target off the front (black point) of an initial set of 100 phenotypes close to archetype 1 in a case of three tasks. The target is equally distant parallel and perpendicular to the front. Each blue point represents the mean phenotype in the population at time steps of one generation of selection towards the target with $p = 0.1$. The population rapidly evolves to the projection of the target on the plane of the triangle, followed by much slower evolution off the triangle. Inset: Parallel (red) and perpendicular (blue) distance from the initial population mean as a function of generation number. (*b*) Response to selection towards a distant target along front (red) and perpendicular to the front (blue). Each point represents the mean phenotype at subsequent generations of selection. Initial population was evolved for 10 000 generations with $N = 100$, $\mu = 0.5$, $k = 2$, free-recombination, before artificial selection for another 10 generations. Artificial selection used $p = 0.5$. (*c*) Artificial selection experiments by Allen *et al.* [31] on butterfly eyespot-size show that the response along the main axis of phenotypic variation of the natural population is larger than the response to selection perpendicular to this axis. The axes are the normalized sizes of the two dorsal eye-spots, and the mean phenotype of the initial population is at the origin. Adapted from [31]. (*d*) Extended simulations of artificial selection perpendicular to the Pareto front results in a line parallel to the front. Initial population was same as in (*b*), artificial selection was performed for 100 generations. (*e*) Experiments by Conner on radish artificial selection perpendicular to the axis of natural variation result in mean offspring on a line (red, blue) parallel to the axis of natural variation (green). Adapted from [32]. Radish image is modified from https://commons.wikimedia.org/wiki/File:Wild_Radish_flower_(5360402586).jpg. (*f*) Normalized beak dimensions of Darwin's ground finch species fall approximately on a plane in the space of beak traits. Perturbations of the beak morphogen pathways, BMP, calmodulin and the pre-maxillary-bone modification factors denoted B, C and P in the figure, have phenotypic effects (arrows) approximately aligned with this plane (generated using data from [33]). See electronic supplementary material, §16 for details. Finch image is modified from https://commons.wikimedia.org/wiki/File:Sparrow_Beak.jpg (*g*) Phalanx proportions of different species of birds fall in a triangle. Populations of chickens and zebra-finches show variation aligned with this triangle. Adapted from [11]. Toe image is modified from [11].

structure. We find in simulations with linkage that polymorphisms still tend to align with the front. However, alignment at a given population size is weaker than in the case of free recombination. Alignment with the front even occurs in simulations with asexual reproduction (no recombination). This is relevant to micro-organisms, and perhaps to tumours, whose

gene expression has been suggested to show trade-offs between tasks [16].

To demonstrate a mechanistic model of trade-offs and gene–gene epistasis, we analysed a molecular mechanism that can give rise to polytope-shaped fronts, based on bacterial sigma-factor regulation. Here, k regulators compete over a limiting factor (e.g. RNA polymerase). This competition naturally gives rise to a polytope with k vertices in gene expression space. Each vertex corresponds to the expression if there was only one sigma factor present, allowing specialization in a certain task. This mechanism can also provide phenotypic plasticity, the ability of an organism to change its phenotype in response to changes in the environment. Plasticity occurs in this mechanism when environmental signals affect the levels of the regulators through upstream pathways—just as sigma factor activities in *E. coli* are affected by stresses and nutrients. The resulting plasticity moves the phenotypes along the Pareto front and not off of it, reaching optimal solutions in different environments.

The Pareto front in empirical examples from morphology has straight edges when using traits that are traditionally used by morphologists such as bone and tooth areas. However, if these traits are nonlinearly transformed, e.g. from T_i to $\sqrt{T_i}$, the fronts change from straight to curved. Thus, plotting the data with traits of bone volume or length instead of area would lead to curved fronts. Curved fronts are more difficult for polymorphisms to align to than straight fronts. This raises an interesting hypothesis: the effects of common mutant alleles should be additive in the traits that provide straight fronts (e.g. additive in effects on bone/tooth area rather than volume or length). In this way, additive effects can best keep offspring along the front. In the case of gene expression, in which straight fronts have been observed in log-transformed data [16,17], this prediction means that mutation effects should be additive in log space (multiplicative effects). Indeed, mutation effects often combine in a multiplicative fashion [39,40].

Other ways that phenotypes can evolve to be on a line have been suggested. These mechanisms do not involve multiple tasks, but instead involve migration [41] or consider a fitness landscape with a pronounced ridge [6,7]. These studies do not directly apply to the present case which employs selection in multiple locales with different conditions. More generally,

rather than assuming the existence of a fitness ridge *a priori*, the Pareto picture suggests a natural explanation for geometries such as lines, triangles and higher-order polytopes in trait space based on the positions of the archetypes that emerge from evolutionary trade-off between the tasks.

The alignment of polymorphisms with the Pareto front has several implications. First, the progeny of any two parents in the population is likely to be on the Pareto front, as long as the mutations are additive or at least have epistatic effects parallel to the front, such as the epistatic effects described here. Without polymorphism alignment, mating would often result in phenotypes off of the front, which would be at a disadvantage because there exists a potential phenotype on the front which is better at all tasks.

A second feature of the alignment of polymorphisms with the Pareto front is rapid response to selective pressures along the front. Evolutionary response is accelerated towards new selection pressures provided that they relate to the same underlying tasks—and therefore align with the Pareto front. Evidence for rapid response to selection pressures along the main axis of phenotypic variation has been reviewed [42] (figure 5c).

To further test the present conclusions requires data on many phenotypic traits together with genomic information for a large number of individuals. Such datasets are expected to become more prevalent in the near future [43,44]. It will be fascinating to explore to what extent the effects of common polymorphisms align with the Pareto front of phenotypic traits.

References

1. Kenney-Hunt JP *et al.* 2008 Pleiotropic patterns of quantitative trait loci for 70 murine skeletal traits. *Genetics* **178**, 2275–2288. (doi:10.1534/genetics.107.084434)

2. Solovieff N, Cotsapas C, Lee PH, Purcell SM, Smoller JW. 2013 Pleiotropy in complex traits: challenges and strategies. *Nat. Rev. Genet.* **14**, 483–495. (doi:10.1038/nrg3461)

3. Johnson T, Barton N. 2005 Theoretical models of selection and mutation on quantitative traits. *Phil. Trans. R. Soc. B* **360**, 1411–1425. (doi:10.1098/rstb.2005.1667)

4. Shoval O, Sheftel H, Shinar G, Hart Y, Ramote O, Mayo A, Dekel E, Kavanagh K, Alon U. 2012 Evolutionary trade-offs, pareto optimality, and the geometry of phenotype space. *Science* **336**, 1157–1160. (doi:10.1126/science.1217405)

5. Smith JM. 1998 *Evolutionary genetics.* Oxford, UK: Oxford University Press.

6. Jones AG, Arnold SJ, Bürger R. 2003 Stability of the G-matrix in a population experiencing pleiotropic mutation, stabilizing selection, and genetic drift. *Evol. Int. J. Org. Evol.* **57**, 1747–1760. (doi:10.1111/j.0014-3820.2003.tb00583.x)

7. Jones AG, Arnold SJ, Bürger R. 2007 The mutation matrix and the evolution of evolvability. *Evolution* **61**, 727–745. (doi:10.1111/j.1558-5646.2007.00071.x)

8. Bowman RI. 1961 *Morphological differentiation and adaptation in the Galapagos finches*, vol. 58. Berkeley, CA: Univ. Calif. Press.

9. Arnold SJ. 1983 Morphology, performance and fitness. *Am. Zool.* **23**, 347–361. (doi:10.1093/icb/23.2.347)

10. Sheftel H, Shoval O, Mayo A, Alon U. 2013 The geometry of the Pareto front in biological phenotype space. *Ecol. Evol.* **3**, 1471–1483. (doi:10.1002/ece3.528)

11. Kavanagh KD, Shoval O, Winslow BB, Alon U, Leary BP, Kan A, Tabin CJ. 2013 Developmental bias in the evolution of phalanges. *Proc. Natl Acad. Sci. USA* **110**, 18190–18195. (doi:10.1073/pnas.1315213110)

12. Tendler A, Mayo A, Alon U. 2015 Evolutionary tradeoffs, Pareto optimality and the morphology of ammonite shells. *BMC Syst. Biol.* **9**, 12. (doi:10.1186/s12918-015-0149-z)

10

rstb.royalsocietypublishing.org   *Phil. Trans. R. Soc. B* **373**: 20170105

13. Szekely P, Korem Y, Moran U, Mayo A, Alon U. 2015 The mass-longevity triangle: pareto optimality and the geometry of life-history trait space. *PLoS Comput. Biol.* **11**, e1004524. (doi:10.1371/journal.pcbi.1004524)

14. Gallagher T, Bjorness T, Greene R, You Y-J, Avery L. 2013 The geometry of locomotive behavioral states in *C. elegans*. *PLoS ONE* **8**, e59865. (doi:10.1371/journal.pone.0059865)

15. Jordan D, Kuehn S, Katifori E, Leibler S. 2013 Behavioral diversity in microbes and low-dimensional phenotypic spaces. *Proc. Natl. Acad. Sci. USA* **110**, 14 018 – 14 023. (doi:10.1073/pnas.1308282110)

16. Hart Y, Sheftel H, Hausser J, Szekely P, Ben-Moshe NB, Korem Y, Tendler A, Mayo AE, Alon U. 2015 Inferring biological tasks using Pareto analysis of high-dimensional data. *Nat. Methods* **12**, 233 – 235. (doi:10.1038/nmeth.3254)

17. Korem Y, Szekely P, Hart Y, Sheftel H, Hausser J, Mayo A, Rothenberg ME, Kalisky T, Alon U. 2015 Geometry of the gene expression space of individual cells. *PLoS Comput. Biol.* **11**, e1004224. (doi:10.1371/journal.pcbi.1004224)

18. Thøgersen JC, Mørup M, Damkiær S, Molin S, Jelsbak L. 2013 Archetypal analysis of diverse *Pseudomonas aeruginosa* transcriptomes reveals adaptation in cystic fibrosis airways. *BMC Bioinformatics* **14**, 279. (doi:10.1186/1471-2105-14-279)

19. Norberg UM, Rayner JMV. 1987 Ecological morphology and flight in bats (Mammalia; Chiroptera): wing adaptations, flight performance, foraging strategy and echolocation. *Phil. Trans. R. Soc. B* **316**, 335 – 427. (doi:10.1098/rstb.1987.0030)

20. Kimura M. 1969 The number of heterozygous nucleotide sites maintained in a finite population due to steady flux of mutations. *Genetics* **61**, 893 – 903.

21. Watterson GA. 1975 On the number of segregating sites in genetical models without recombination. *Theor. Popul. Biol.* **7**, 256 – 276. (doi:10.1016/0040-5809(75)90020-9)

22. Pond SLK, Posada D, Gravenor MB, Woelk CH, Frost SDW. 2006 Automated phylogenetic detection of recombination using a genetic algorithm. *Mol. Biol. Evol.* **23**, 1891 – 1901. (doi:10.1093/molbev/msl051)

23. Falconer DS, Mackay TFC. 1996 *Introduction to quantitative genetics*, 4th edn. Essex, England: Pearson.

24. Gillespie JH. 2004 *Population genetics: A concise guide*, 2nd edn. Baltimore, MD: Johns Hopkins University Press.

25. Lynch M, Walsh B. 1998 *Genetics and analysis of quantitative traits*, 1st edn. Sunderland, MA: Sinauer Associates.

26. Hengge R. 2011 The general stress response in gram-negative bacteria. In *Bacterial stress responses, second edition* (eds R Hengge, G Storz), pp. 251 – 289. American Society of Microbiology.

27. Nyström T. 2004 MicroReview: growth versus maintenance: a trade-off dictated by RNA polymerase availability and sigma factor competition? *Mol. Microbiol.* **54**, 855 – 862. (doi:10.1111/j.1365-2958.2004.04342.x)

28. Dekel E, Alon U. 2005 Optimality and evolutionary tuning of the expression level of a protein. *Nature* **436**, 588 – 592. (doi:10.1038/nature03842)

29. Kafri M, Metzl-Raz E, Jona G, Barkai N. 2016 The cost of protein production. *Cell Rep.* **14**, 22 – 31. (doi:10.1016/j.celrep.2015.12.015)

30. Jones AG, Bürger R, Arnold SJ. 2014 Epistasis and natural selection shape the mutational architecture of complex traits. *Nat. Commun.* **5**, 3709. (doi:10.1038/ncomms4709)

31. Allen CE, Beldade P, Zwaan BJ, Brakefield PM. 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)

32. Conner JK. 2003 Artificial selection: a powerful tool for ecologists. *Ecology* **84**, 1650 – 1660. (doi:10.1890/0012-9658(2003)084[1650:ASAPTF]2.0.CO;2)

33. Mallarino R, Grant PR, Grant BR, Herrel A, Kuo WP, Abzhanov A. 2011 Two developmental modules establish 3D beak-shape variation in Darwin's finches. *Proc. Natl. Acad. Sci. USA* **108**, 4057 – 4062. (doi:10.1073/pnas.1011480108)

34. Agrawal AA, Conner JA, Rasmann S. 2010 Tradeoffs and negative correlations in evolutionary ecology. In *Evolution since Darwin: the first 150* (eds MA Bell, DJ Futuyma, WF Eanes, JS Levinton), pp. 243 – 268. Sunderland, MA: Sinauer.

35. Abzhanov A. 2004 Bmp4 and morphological variation of beaks in Darwin's finches. *Science* **305**, 1462 – 1465. (doi:10.1126/science.1098095)

36. Abzhanov A, Kuo WP, Hartmann C, Grant BR, Grant PR, Tabin CJ. 2006 The calmodulin pathway and evolution of elongated beak morphology in Darwin's finches. *Nature* **442**, 563 – 567. (doi:10.1038/nature04843)

37. Grant PR, Abbott I, Schluter D, Curry RL, Abbott LK. 1985 Variation in the size and shape of Darwin's finches. *Biol. J. Linn. Soc.* **25**, 1 – 39. (doi:10.1111/j.1095-8312.1985.tb00384.x)

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

39. Velenich A, Gore J. 2013 The strength of genetic interactions scales weakly with mutational effects. *Genome Biol.* **14**, R76. (doi:10.1186/gb-2013-14-7-r76)

40. Friedlander T, Mayo AE, Tlusty T, Alon U. 2013 Mutation rules and the evolution of sparseness and modularity in biological systems. *PLoS ONE* **8**, e70444. (doi:10.1371/journal.pone.0070444)

41. Guillaume F, Whitlock MC. 2007 Effects of migration on the genetic covariance matrix. *Evolution* **61**, 2398 – 2409. (doi:10.1111/j.1558-5646.2007.00193.x)

42. Klingenberg CP. 2010 Evolution and development of shape: integrating quantitative approaches. *Nat. Rev. Genet.* **11**, 623 – 635. (doi:10.1038/nrg2829)

43. O'Reilly PF, Hoggart CJ, Pomyen Y, Calboli FCF, Elliott P, Jarvelin M-R, Coin LJM. 2012 MultiPhen: joint model of multiple phenotypes can increase discovery in GWAS. *PLoS ONE* **7**, e34861. (doi:10.1371/journal.pone.0034861)

44. Bush WS, Oetjens MT, Crawford DC. 2016 Unravelling the human genome-phenome relationship using phenome-wide association studies. *Nat. Rev. Genet.* **17**, 129 – 145. (doi:10.1038/nrg.2015.36)