

Epistasis from functional dependence of fitness on underlying traits

Hsuan-Chao Chiu, Christopher J. Marx and Daniel Segrè

Proc. R. Soc. B published online 15 August 2012
doi: 10.1098/rspb.2012.1449

Supplementary data

"Data Supplement"

<http://rsob.royalsocietypublishing.org/content/suppl/2012/08/08/rspb.2012.1449.DC1.html>

References

This article cites 46 articles, 17 of which can be accessed free

<http://rsob.royalsocietypublishing.org/content/early/2012/08/08/rspb.2012.1449.full.html#ref-list-1>

P<P

Published online 15 August 2012 in advance of the print journal.

EXiS Open Choice

This article is free to access

Subject collections

Articles on similar topics can be found in the following collections

[evolution](#) (1260 articles)
[systems biology](#) (29 articles)
[theoretical biology](#) (34 articles)

Email alerting service

Receive free email alerts when new articles cite this article - sign up in the box at the top right-hand corner of the article or click [here](#)

Advance online articles have been peer reviewed and accepted for publication but have not yet appeared in the paper journal (edited, typeset versions may be posted when available prior to final publication). Advance online articles are citable and establish publication priority; they are indexed by PubMed from initial publication. Citations to Advance online articles must include the digital object identifier (DOIs) and date of initial publication.

Epistasis from functional dependence of fitness on underlying traits

Hsuan-Chao Chiu^{1,†}, Christopher J. Marx³ and Daniel Segre^{1,2,*}

¹*Bioinformatics Program, and* ²*Department of Biology and Department of Biomedical Engineering, Boston University, Boston, MA 02215, USA*

³*Department of Organismic and Evolutionary Biology and Faculty of Arts and Sciences Center for Systems Biology, Harvard University, Cambridge, MA 02138, USA*

Epistasis between mutations in two genes is thought to reflect an interdependence of their functions. While sometimes epistasis is predictable using mechanistic models, its roots seem, in general, hidden in the complex architecture of biological networks. Here, we ask how epistasis can be quantified based on the mathematical dependence of a system-level trait (e.g. fitness) on lower-level traits (e.g. molecular or cellular properties). We first focus on a model in which fitness is the difference between a benefit and a cost trait, both pleiotropically affected by mutations. We show that despite its simplicity, this model can be used to analytically predict certain properties of the ensuing distribution of epistasis, such as a global negative bias, resulting in antagonism between beneficial mutations, and synergism between deleterious ones. We next extend these ideas to derive a general expression for epistasis given an arbitrary functional dependence of fitness on other traits. This expression demonstrates how epistasis relative to fitness can emerge despite the absence of epistasis relative to lower level traits, leading to a formalization of the concept of independence between biological processes. Our results suggest that epistasis may be largely shaped by the pervasiveness of pleiotropic effects and modular organization in biological networks.

Keywords: epistasis; benefit–cost model; evolution; pleiotropy; modularity

1. INTRODUCTION

Epistasis describes a fundamental nonlinearity in biological systems, capturing the fact that the phenotypic effect of a genetic mutation or allele variant could depend on another mutation. Epistasis is suggested to play an important role in evolutionary dynamics, e.g. by shaping the fitness landscape [1,2], maintaining sexual reproduction [3,4] and affecting the speed of adaptation [5–8]. Large-scale systematic studies of single and double gene deletions have also revealed that epistasis is useful in understanding the organization of living systems into modules, as genes belonging to the same biological process tend to share similar profiles of epistatic interactions with other genes [9,10]. Given the influence of epistasis on evolutionary processes, and given the modular structure of epistatic interaction networks, one may wonder whether deeper insights can be obtained on how the modular organization of biological networks affects (and is affected by) epistasis and adaptation. The connection between modularity and evolution has been addressed from multiple independent standpoints [11–16]. Yet few concepts seem to be more appropriate than epistasis for trying to relate the architecture of biological networks with the evolutionary forces that gave rise to such networks. This is ultimately due to the fact that epistasis captures in an elementary way the complexity of the genotype–phenotype map.

For quantitative traits, epistasis can be analysed mathematically. Upon a double mutation, a given quantitative trait could turn out to be higher or lower than expected based on individual mutations, leading to a quantifiable positive or a negative epistasis, respectively. Special caution should be used in comparing the definitions of epistasis across different research areas, as multiple, sometime conflicting, terminologies have been adopted. A commonly used classification is the one between synergistic and antagonistic epistasis. Synergistic (antagonistic) epistasis occurs when the joint effect of two alleles is more (less) severe than expected. Note that, based on the terminology adopted here, negative epistasis between deleterious mutations corresponds to synergism, while negative epistasis between beneficial mutations is classified as antagonism (table 1). In addition, the very definition of the baseline expectation relative to which epistasis is quantified is still the subject of active debate [17–20]. In this work, we assume that mutational effects combine multiplicatively in the absence of epistasis (see also §2).

The role of epistasis in evolution constitutes a particularly important and debated question [21], which involves understanding the sources and the consequences of the average and variance of the distribution of epistatic effects [22]. Therefore, considerable effort has been put into elucidating the distribution of epistasis, both from a theoretical and from an experimental perspective [23,24]. Some early experimental studies, partially motivated by the possible relevance of synergistic effects between deleterious mutations in the evolution of sex [3,4], had found nearly symmetric distributions of epistasis [25–28]. Subsequent large-scale studies under different proportions of beneficial versus deleterious mutations have reported both negative

* Author for correspondence (dsegre@bu.edu).

† Present address: Department of Genome Sciences, University of Washington, Seattle, WA 98195, USA.

Electronic supplementary material is available at <http://dx.doi.org/10.1098/rsob.2012.1449> or via <http://rsob.royalsocietypublishing.org>.

Table 1. Definitions and conventions for epistasis in the current work. We quantify the degree of epistasis (ε) as the deviation of the joint effect of mutations from the expectation in multiplicative scale (see equation (2.1)). Synergistic epistasis occurs when the joint effect of two alleles is reinforced (e.g. more severe than the multiplicative expectation), while antagonistic epistasis happens when the joint effect is buffered (less severe than the multiplicative expectation) by the interaction between alleles. As illustrated in the table, negative epistasis ($\varepsilon < 0$) may point to synergistic or antagonistic behaviour based on whether the mutations are both beneficial or both deleterious.

	type of mutations	
	deleterious	beneficial
synergistic	$\varepsilon < 0$	$\varepsilon > 0$
antagonistic	$\varepsilon > 0$	$\varepsilon < 0$

[10] and positive [29] epistasis trends, leaving the problem fundamentally unsettled. From a theoretical perspective, models of fitness landscapes have been used for providing potential explanations for observed epistasis trends. Antagonism between beneficial mutations, for example, can be predicted by a fitness landscape model that assumes ubiquity of stabilizing selection [30]. In addition, stoichiometric models of metabolic networks have been useful in exploring the distribution and network organization of gene–gene interactions in metabolism [9,31], and in providing mechanistic explanations for measured distributions [29,32]. However, none of these prior theoretical works seems to have explicitly addressed the question of how epistasis quantitatively depends on the modular organization of the genotype–phenotype mapping, and on the degree of pleiotropy.

A hint to how modularity and epistasis relate to each other was recently offered by an experimental evolution study that identified a diminishing returns trend among pairs of beneficial alleles [7] (see also recent studies [8,33]). This study reported antagonistic epistasis between beneficial alleles that arose during 600 generations of evolution of an engineered strain of *Methylobacterium extorquens*. It was found that the observed beneficial alleles improve fitness either by enhancing metabolic capacity or by alleviating protein expression-related costs. This observation suggested that microbial fitness could very coarsely be treated as a modularly structured function of two separate phenotypic traits, a metabolic benefit and a protein expression cost. In fact, a simple benefit–cost model was sufficient to quantitatively predict fitness values for multi-allele strains, and to quantitatively recapitulate the antagonistic trend for epistasis between genes in different loci of the genome [7]. Benefit–cost models had been used to describe fitness in previous studies of evolutionary adaptation, ranging from rapid adaptation of protein expression in new environments [14], to optimal regulatory design upon biochemical noise [34] and optimal transcriptional regulation of metabolism [35].

Here, inspired by the previously demonstrated relevance of benefit–cost models in evolutionary research, we use a benefit–cost model as a starting point for quantifying epistasis based on the dependence of fitness on multiple traits. In particular, we first extend the

benefit–cost model from a way to explain an observed case of antagonism between beneficial mutations [7], to a general statistical analysis of expected epistasis distributions among mutations with a given chance of being beneficial or deleterious. In the second part of our work, we generalize these results to show that the degree of epistasis can be estimated analytically for an arbitrary dependence of fitness on simpler quantitative traits, providing a new mathematical link between epistasis, pleiotropy and modularity. Note that our analysis is mostly focused on analysing the interactions between two mutations, without delving into the problem of how multiple perturbations accumulate. In addition, we restrict our calculations almost entirely to a fitness function that depends on two traits (also previously called two-dimensional epistasis [36]), leaving possible extensions to multidimensional epistasis [36] as a topic for potential follow-up research.

2. BACKGROUND AND DEFINITIONS

The main premise of the current work is that a high-level trait, or phenotype, such as fitness (f) can be phenomenologically expressed as a function F of two basic observable traits, X and Y : $f = F(X, Y)$. For example, the growth rate of a bacterium may be expressed as a function of its respiratory and fermenting capacities [37]. In addition to a wild-type organism (e.g. a bacterial strain), we consider two mutant strains with genetic modifications at loci i and j , respectively, and a double mutant strain which has both i and j alleles modified. For each of these four strains, we hypothesize that it is possible to independently measure the overall fitness (f_0 for the wild-type, f_i and f_j for the single mutants, and f_{ij} for the double mutant), as well as each of the two basic traits X and Y (x_0 for the wild-type, x_i and x_j for the single mutants, and x_{ij} for the double mutant, etc.). The general question we are concerned with is whether we can estimate epistasis with regard to fitness between perturbations i and j given some assumptions on how these mutations affect phenotypes X and Y , and given the functional dependence of F on X and Y (figure 1).

In the absence of epistasis, it is often assumed that mutational effects should combine multiplicatively, i.e. $f_{ij}/f_0 = (f_i/f_0)(f_j/f_0)$ ([7–10,25,31,38], see also discussion on alternative definitions [39]). In other words, under this assumption, absence of epistasis means that perturbations will tend to affect fitness in a way that is proportional to the fitness itself, and independent of the genetic background (as easily seen by rewriting the previous expression as $f_{ij}/f_j = f_i/f_0$). Epistasis can be therefore seen as a deviation from this null multiplicative expression and quantified as

$$\varepsilon = \frac{f_{ij}}{f_0} - \frac{f_i}{f_0} \cdot \frac{f_j}{f_0}. \quad (2.1)$$

This equation has been used previously [7–10,25,31,38] for assessing the degree of epistasis, given experimentally measured or computationally predicted fitness values for the wild-type and mutant strains. It is through this equation that one can determine whether epistasis is positive ($\varepsilon > 0$) or negative ($\varepsilon < 0$). In this work, we will use this equation to determine how an analytical expression for fitness as a function of other

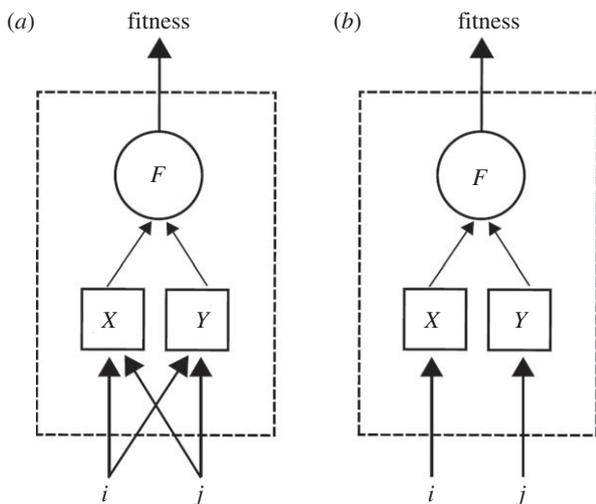


Figure 1. Schematic depiction of how we quantify epistasis relative to a fitness function that depends on two quantitative traits, or phenotypes. (a) Two alleles or genetic perturbations i and j are assumed to potentially affect multiple traits, here X and Y ('low-level traits'). The phenomenon in which a genetic perturbation affects multiple traits is called pleiotropy. Here we assume that there is no epistasis at the level of the individual traits X and Y . A 'high-level trait' (e.g. fitness f) is defined as a function F of the two traits X and Y . These assumptions allow us to predict how the functional shape of F affects epistasis between the two perturbations. Without any knowledge of this internal structure (dashed box), the presence of epistasis could only be measured experimentally, but not inferred mathematically. (b) The same model as described above, in the absence of pleiotropy. In this case, perturbations i and j affect each a single trait, i.e. X and Y respectively, and can be thought of acting on different modules. Depending on the function F , this may still lead to epistasis.

measurable traits (and of the mutations that affect such traits), translates into a prediction of epistasis. We will restrict our analysis to the simplest case in which alleles exhibit no direct epistasis with regard to either of the two underlying traits X and Y (i.e. $x_{ij}/x_0 = (x_i/x_0)(x_j/x_0)$ and $y_{ij}/y_0 = (y_i/y_0)(y_j/y_0)$). Under this assumption, epistasis relative to fitness emerges only as a consequence of the functional dependence of F on X and Y (figure 1).

The benefit–cost model used to explain an antagonistic epistasis pattern that emerged during adaptation of an *M. extorquens* strain [7] constitutes a special case of the F function introduced earlier. In this case, in analogy to Dekel & Alon [14], $X = b$ is the growth advantage (benefit), and $Y = c$ is the burden (cost) associated with the operation of the biological network, and $F = X - Y = b - c$. In this model, fitness of the wild-type strain can be written as the difference between a basal benefit and a basal cost term:

$$f_0 = b_0 - c_0 = 1. \quad (2.2)$$

The quantities f_0 and c_0 can be experimentally measured to operationally define the basal benefit b_0 as demonstrated in earlier studies [7,14]. A mutant allele i is assumed to pleiotropically alter b_0 or c_0 with factors λ_i and θ_i respectively, yielding a fitness:

$$f_i = f(b_i, c_i) = \lambda_i b_0 - \theta_i c_0. \quad (2.3)$$

To avoid confusion, it is important to stress that λ_i represents the effect of a single mutation (i) on the benefit component of the fitness (if i has no effect on the benefit, then $\lambda_i = 1$), and is not proportional to the number of mutations accumulated in a longer-term evolutionary process. Rather, multiple mutations are assumed to act by the action of further multiplicative factors. For example, a successive allele j , on the background of i , is assumed again to act multiplicatively on the benefit and cost components, giving rise to the fitness of the double mutant

$$f_{ij} = f(b_{ij}, c_{ij}) = \lambda_i \lambda_j b_0 - \theta_i \theta_j c_0. \quad (2.4)$$

Note that the changes in each trait are not simply given by the λ and θ factors, but rather also include a dependence on the unperturbed trait, e.g. $\Delta b = \lambda_i b_0 - b_0 = (\lambda_i - 1)b_0$. In our previous work [7], experimentally determined values of b_0 , c_0 , and of λ_i and θ_i for each allele, were used in equation (2.4) (and its extension to more than two mutations) to provide accurate predictions of the fitness of multi-allele strains, and to explain the observed antagonistic epistasis among some beneficial alleles. Here instead, we explore the space of possible pairs of mutational effects to infer statistical properties of epistasis.

3. RESULTS

(a) An analytical expression of epistasis in the benefit–cost model

Our first goal is to determine analytically the magnitude and sign of epistasis relative to fitness under the assumptions of the benefit–cost model. As shown in detail in electronic supplementary material, §B, this can be achieved by substituting the expressions for the fitness of single and double mutants (equations (2.3) and (2.4)) into the definition of epistasis (equation (2.1)), yielding:

$$\varepsilon = -b_0 c_0 (\lambda_i - \theta_i)(\lambda_j - \theta_j). \quad (3.1)$$

This is a remarkably simple expression, in which epistasis turns out to be computable as the product of a term that depends only on the unperturbed state parameters (b_0 , c_0), and on a term that depends only on the phenotypic effects of the mutations (the λ_i , θ_i , λ_j and θ_j parameters). From equation (3.1), one can see that no epistasis ensues in the benefit–cost model if $\lambda = \theta$ for one or both mutations (see electronic supplementary material, figure S9 for additional clarifications, including an extension to more than two mutations). Importantly, given that the benefit and cost terms are defined as positive ($b_0 > 1$ and $c_0 > 0$), the sign of epistasis is entirely determined by the $(\lambda_i - \theta_i)(\lambda_j - \theta_j)$ product. Hence, in this model, the sign of epistasis between two mutations depends only on their mutational effects on the benefit and the cost, but not on the initial unperturbed values. In addition, from equation (3.1), one can see that positive epistasis can be obtained only by combining mutations that have $(\lambda - \theta)$ values of opposite sign. Additional steps are required to determine how this expression for epistasis (and especially its sign) depends on whether individual mutations are beneficial or deleterious.

(b) The benefit–cost model imposes a negative bias in the distribution of epistasis

Equation (3.1) predicts the degree of epistasis for specific combinations of two mutations. In this section, we show

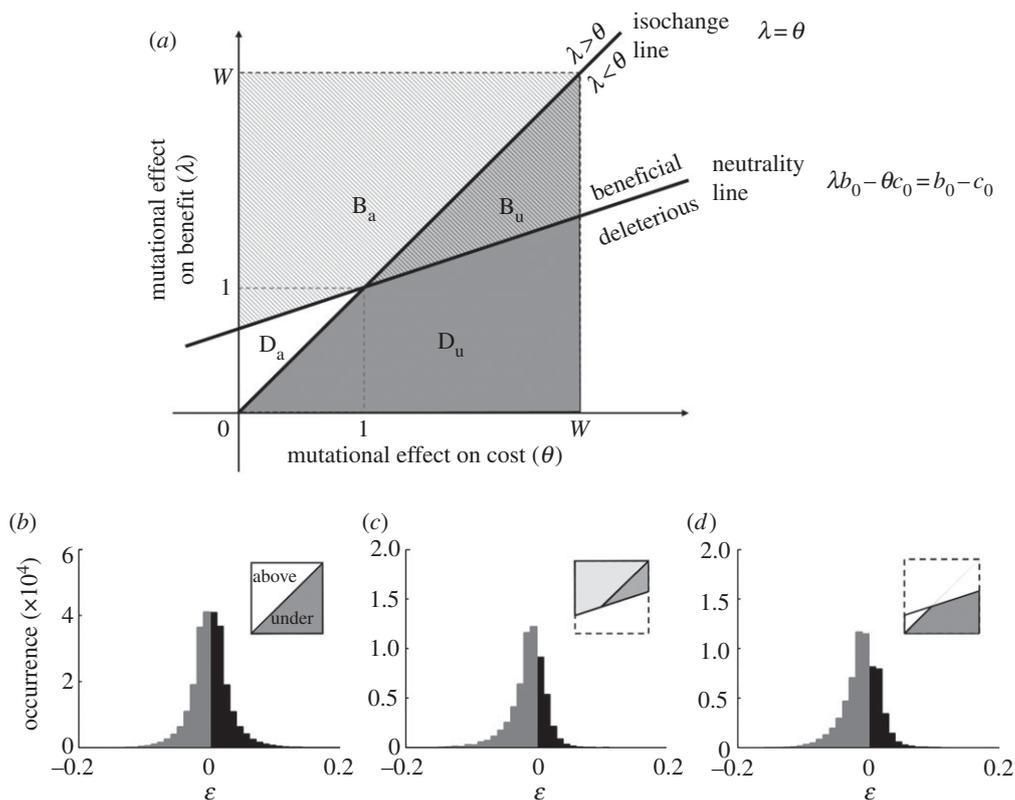


Figure 2. Estimating epistasis through a geometrical representation of perturbations in phenotype space. (a) The (λ, θ) plane, a geometrical representation of possible mutant alleles in a benefit–cost model of fitness. Any allele (e.g. i) can be represented as a point with coordinates (λ_i, θ_i) corresponding to the multiplicative alterations of the benefit and cost, respectively. We assume that both λ and θ can have values between zero and W . Throughout the paper, we assume $W = 2$, so that beneficial and deleterious mutations have equal chance of being chosen when sampling uniformly. The (λ, θ) plane is divided into four regions by the neutrality line (corresponding to mutants with fitness equal to the wild-type) and the isochange line (corresponding to mutations such that $\lambda_i = \theta_i$). The intersection between these two lines (i.e. the point $(\lambda, \theta) = (1, 1)$) corresponds to the wild-type strain. B_a is the area containing beneficial alleles above the isochange line; B_u is the area containing beneficial alleles under the isochange line. D_a and D_u are similarly defined for deleterious alleles. The combination of two alleles both lying above the isochange line will give rise to negative ε , as evident from equation (3.1). In general, the sign of ε depends on the chance of selecting alleles from different regions in the (λ, θ) plane. The maximum value of $B_u = (W - 1)^2/2$ occurs when the slope of the neutrality line is zero ($c_0 = 0$). The corresponding B_a in this situation is $B_u + (W - 1)$. When we increase the slope, B_u decreases (while B_a increases) monotonically as c_0 goes up, until B_u reaches its minimum value at zero when $c_0 = b_0$ (slope of neutrality line = 1). Thus, it is always $B_a > B_u$. (b) Without imposing any constraint on whether mutations are beneficial or deleterious the regions above and under the isochange line have equal chance to occur (inset), leading to an unbiased epistasis distribution. (c,d) Negative bias between strictly beneficial alleles (c, region $B_a > B_u$ shaded in inset) and between strictly deleterious alleles (d, region $D_u > D_a$ shaded in inset) can be demonstrated analytically, and is confirmed here by simulations (see the electronic supplementary material, §A).

how this same equation can be also used to draw general conclusions about the expected distribution of epistasis between a pair of perturbations in the benefit–cost model. The fact that the sign of epistasis is dictated by the product of $(\lambda - \theta)$ terms suggests that a statistical analysis of how these terms are distributed could provide information about the distribution of epistasis itself.

To reason about this problem, one can visualize possible choices of perturbations on the (λ, θ) plane (figure 2). Each perturbation in the benefit–cost model can be represented as a point in this plane. The point of coordinates $(1, 1)$ corresponds to the wild-type, i.e. the unperturbed system. For simplicity, we assume here that λ and θ cannot exceed a given value W . Two fundamental lines can be drawn on this plane. One line, which we call the *isochange line*, is defined by the equation $\lambda = \theta$, and corresponds to all individual perturbations that change both the benefit and the cost by the same multiplicative factor. Note that the isochange line does not have in itself an immediate interpretation in terms of epistasis (e.g. it

does not represent a boundary between positive and negative epistasis); rather, as reasoned later, it is an abstract geometrical construct that will help us to determine the chance of observing a given sign of epistasis for two mutations, through equation (3.1). If we think of two mutations as two points with coordinates (λ_i, θ_i) and (λ_j, θ_j) , the positions of these points relative to the *isochange line* will determine the signs of the terms $(\lambda_i - \theta_i)$ and $(\lambda_j - \theta_j)$, and hence, based on equation (3.1), the sign of epistasis between such mutations. If no other constraints exist in the system, upon uniformly sampling pairs of points in the (λ, θ) plane (with $W = 2$), it is equally likely to choose positive or negative $(\lambda - \theta)$ terms, giving rise to no obvious bias in the distribution of ε (figure 2b).

A second fundamental line in the (λ, θ) plane is the line that partitions beneficial from deleterious mutations. This line, which we call *neutrality line* ($\lambda_i b_0 - \theta_i c_0 = 1$), corresponds to all possible choices of λ and θ whose combined effect is to leave fitness equal to the wild-type value of 1. All points above this line are associated with

Table 2. Contingency table for the phenotypic values of strictly beneficial alleles. The categories of ε classified by the four conditions are analogous to the possible outcomes of tossing a coin twice, and allow us to compute the overall probability of negative and of positive ε , giving $p(\varepsilon < 0) - p(\varepsilon > 0) = (B_a^2 + B_u^2 - 2B_aB_u)/B_{\text{tot}}^2 = (B_a - B_u)^2/B_{\text{tot}}^2 \geq 0$, where $B_{\text{tot}} = B_a + B_u$.

ε	-	-	+	+
condition	$\lambda_i > \theta_i$	$\lambda_i < \theta_i$	$\lambda_i > \theta_i$	$\lambda_i < \theta_i$
	$\lambda_j > \theta_j$	$\lambda_j < \theta_j$	$\lambda_j < \theta_j$	$\lambda_j > \theta_j$
region (i, j)	(B_a, B_a)	(B_u, B_u)	(B_a, B_u)	(B_u, B_a)
$p(\text{condition})$	B_a^2/B_{tot}^2	B_u^2/B_{tot}^2	B_aB_u/B_{tot}^2	B_uB_a/B_{tot}^2
	$p(\varepsilon < 0) = \frac{B_a^2 + B_u^2}{(B_a + B_u)^2}$		$p(\varepsilon > 0) = \frac{2B_aB_u}{(B_a + B_u)^2}$	

individual beneficial mutations; all points below this line correspond to individual deleterious mutations. It is now possible to ask whether any bias in the distribution of ε may be expected among strictly beneficial mutations. In other words, we ask whether a repeated uniform sampling of pairs of points from the region above the neutrality line will preferentially yield positive or negative values of ε , based on equation (3.1). Note that the sampling we are performing here is a sampling in phenotype space, and is not meant to provide insights into the actual rates of beneficial/deleterious mutations in the genome. What determines the sign of ε , for a given pair of beneficial perturbations, is whether points selected above the neutrality line fall above (area B_a) or under (area B_u) the *isochange line*. Specifically, as explained in table 2, it is easy to see that the difference in the chance to obtain a negative versus a positive ε is $p(\varepsilon < 0) - p(\varepsilon > 0) = (B_a - B_u)^2/(B_a + B_u)^2$. This expression is always positive as long as $B_a \neq B_u$, a condition geometrically confirmed by figure 2. Thus, under the benefit–cost model, one should expect a bias towards negative ε (antagonistic epistasis) among beneficial mutations (figure 2c), consistent with the specific observations of [7]. As shown in figure 2d, a similar result can be obtained for deleterious mutations (points below the neutrality line). In this case, again, one can infer a tendency towards negative ε , indicating a bias towards synergistic epistasis between deleterious mutations. It is possible to further generalize the earlier-mentioned results to perturbations with any given probability ρ of being beneficial (i.e. the fraction of beneficial mutations ρ can be any value other than 0.5). It can be shown (see the electronic supplementary material, §C) that in this general case, one can analytically compute the excess probability of negative epistasis ($\Delta p^\varepsilon = p(\varepsilon < 0) - p(\varepsilon > 0)$) as a function of ρ and c_0 . In particular, under simplifying assumptions, one obtains $\Delta p^\varepsilon = (2\rho - 1)^2 \times (2c_0 + 1)^2 / (2c_0 + 2)^2$ (see the electronic supplementary material, figure S2b). Unless $\rho = 0.5$, this expression is expected to be always positive, demonstrating that the negative bias in the distribution of epistasis is a general property of the benefit–cost model. We further support this analytical conclusion with sensitivity analyses against choices of ρ and c_0 . This negative bias is consistently obtained, with no substantial difference computationally (see the electronic supplementary material, figure S2a) or analytically (see the electronic supplementary material, figure S2b). In addition, no significant deviations from this trend are observed upon introducing an arbitrary interdependence between λ and θ in the form of a rotated

Gaussian bivariate distribution (see the electronic supplementary material, figure S3).

While the geometrical arguments illustrated earlier provide estimates of the bias in the distribution of ε as a function of the fraction, ρ , of beneficial mutations, they do not allow us to predict the shape of distribution itself. We used computer simulations (see the electronic supplementary material, §A) to explore the full spectrum of epistasis distribution as a function of ρ . Again, here ρ is a phenotypic measure; we make no assumption on the connection between random mutations and specific values of ρ . The simulations confirm that negative ε is more likely to occur over the whole range of ρ values (figure 3a). The bias (prevailing negative epistasis) reaches a maximum in both extreme cases (strictly beneficial or strictly deleterious; figure 2c,d) while it becomes less and less pronounced as ρ approaches 0.5 (figure 3a,b). Note that for distributions derived from combinations of beneficial and deleterious mutations, a negative ε cannot be easily associated with synergistic or antagonistic trends, as different pairs contributing to the distribution will have different effects (including sign epistasis, an interesting case where an allele is beneficial on some genetic backgrounds but deleterious on others; see electronic supplementary material, figure S7 for more details). Computer simulations also indicate that these epistasis trends are robust over a broad range of values for c_0 (see the electronic supplementary material, figure S2), and that they would equally ensue in a more complex model involving multiple cost components (e.g. $f_0 = b_0 - c_0 - d_0 - e_0$, electronic supplementary material, figure S1).

(c) Analytical estimate of epistasis for an arbitrary function F

Most of our study so far has been focused on a specific dependence of fitness on two traits, i.e. the difference between a benefit and a cost trait. We next generalize our analysis to ask whether it is possible to estimate epistasis when fitness depends in an arbitrary way on two quantitative traits X and Y , e.g. it is the sum, the product or any arbitrary function F of such traits. As performed for the derivation of equation (3.1), we need to substitute the expressions for fitness of single and double mutant strains into equation (2.1) to compute $\varepsilon = F(x_{ij}, y_{ij})/F(x_{0j}, y_{0j}) - [F(x_{ip}, y_{ip})/F(x_{0p}, y_{0p})] \cdot [F(x_{jp}, y_{jp})/F(x_{0j}, y_{0j})]$. In this case, however, no further result can be obtained unless additional simplifying assumptions are made about the system. One possible such assumption is that mutations cause small

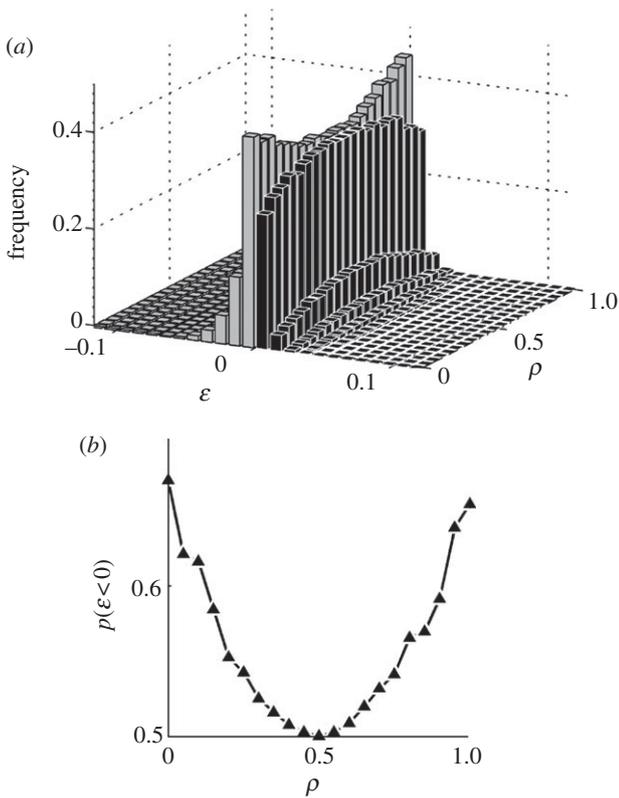


Figure 3. Numerically computed epistasis distributions show a generic negative trend for all possible proportions of beneficial mutations. Each bell-shaped histogram corresponds to the distribution of epistasis at a given fraction of beneficial mutations (ρ). For visual clarity, bars associated with negative ε are depicted in light grey, while bars for positive ε are depicted in dark grey. The front slice ($\rho = 0$) is the same distribution shown in figure 2*d*. (a) The concave shape for the negative ε bars across different values of ρ indicates that the bias towards negative ε increases as the portion of beneficial allele moves away from 0.5. (b) Negative epistasis is more likely to occur when the single mutants are dominated by mostly beneficial ($\rho \gg 0.5$) or mostly deleterious alleles ($\rho \ll 0.5$).

perturbations to the traits ($\Delta x_i = x_i - x_0 \ll x_0$, $\Delta x_j = x_j - x_0 \ll x_0$, $\Delta x_{ij} = x_{ij} - x_0 \ll x_0$, and similarly for Δy). Under this assumption, one can perform a Taylor expansion of each term in the above expression for ε , e.g. for $F(x_i, y_i) = F(x_0 + \Delta x_i, y_0 + \Delta y_i)$. Note that because the product $F(x_i, y_i) \cdot F(x_j, y_j)$ will give rise to second order terms (e.g. in Δx^2), it is essential to perform the Taylor expansion to the second order. In the derivation, presented in detail in electronic supplementary material, §D, we assumed (as for the benefit–cost model) that there is no epistasis between perturbations relative to each of the two traits X and Y . After some algebraic rearrangements, we obtained for epistasis

$$\begin{aligned} \varepsilon = & \Delta x_i \Delta x_j (\dot{F}_X/x_0 + \ddot{F}_{XX} - \dot{F}_X^2) \\ & + \Delta y_i \Delta y_j (\dot{F}_Y/y_0 + \ddot{F}_{YY} - \dot{F}_Y^2) \\ & + (\Delta x_i \Delta y_j + \Delta x_j \Delta y_i) \cdot (\ddot{F}_{XY} - \dot{F}_X \dot{F}_Y), \end{aligned} \quad (3.2)$$

where \dot{F}_X , \dot{F}_Y , \ddot{F}_{XY} , \ddot{F}_{XX} and \ddot{F}_{YY} are the partial derivatives of F computed at x_0, y_0 . The first two terms in (3.2) quantify the contribution to epistasis through individual traits X and Y , respectively, whereas the third term is

associated with their functional coupling. As a generalized form of equation (3.1), equation (3.2) provides a direct quantitative link between the magnitude and direction of epistasis and the magnitude of individual mutations. For a function $F = X - Y$, upon expressing the Δx and Δy parameters in terms of λ and θ , it is easy to show that equation (3.2) yields the benefit–cost result of equation (3.1) (see electronic supplementary material, §D). One can further use equation (3.2) to explore expected epistasis under other possible functional dependencies. For example, if F is a linear combination of the two phenotypes, $F(X, Y) = aX + bY$, then the degree of epistasis converges to a generalized form of equation (3.1), where $\varepsilon = abx_0y_0(\lambda_i - \theta_i)(\lambda_j - \theta_j)$. In this case, the relative signs of a and b determine whether the expected distribution of epistasis will display a positive or negative bias. In addition, as expected, one can verify that, if $F(X, Y) = XY$, then epistasis is always zero, compatible with the idea that multiplicative effects on individual traits will combine to provide an overall multiplicative effect on fitness. Interestingly, this is also true for functions of the form $F(X, Y) = X^n Y^m$, hinting to a broader view of the relationship between epistasis and independence, as explored in detail later.

An important question one can ask using the expression of equation (3.2) is what degree of epistasis should be expected between two mutations each affecting only one of the two traits (i.e. non-pleiotropic mutations; figure 1*b*). Does our model support the general intuition that mutations affecting independent modules in a biological system should have no epistasis? Equation (3.2) allows us to ask this question in a formal way. In our framework, lack of pleiotropy is expressed by assuming, for example, that allele i affects only trait X , and allele j affects only Y (i.e. $\Delta x_j = \Delta y_i = 0$, equivalent to $\lambda_j = \theta_i = 1$, in the benefit–cost model; figure 1*b*). From equation (3.2), one can see that under these conditions epistasis can be expressed as

$$\begin{aligned} \varepsilon = & \Delta x_i \Delta y_j (\ddot{F}_{XY} - \dot{F}_X \dot{F}_Y) \\ = & \frac{\Delta x_i \Delta y_j}{x_0 y_0} \left(\frac{\partial^2 \log F}{\partial X \partial Y} \right) \Big|_{x_0, y_0}. \end{aligned} \quad (3.3)$$

Equation (3.3) implies that, even if each mutation affects only one of the two phenotypes (figure 1*b*), epistasis will be zero only for the set of functions F that satisfy the condition $\partial^2 \log F / \partial x \partial y = 0$. One can verify (see the electronic supplementary material, §D) that this condition is satisfied by any function decomposable as

$$F(X, Y) = G(X)H(Y) \quad (3.4)$$

Note that, as opposed to functions such as $F(X, Y) = X^n Y^m$, which would always give zero epistasis (irrespective of whether mutations are pleiotropic or not), one can think of several other F functions that would satisfy equation (3.4), but give zero epistasis only in the absence of pleiotropy (e.g. $F(X, Y) = \exp(X) \exp(Y)$, see electronic supplementary material, §D).

The consequences of equations (3.2)–(3.4) can be summarized as follows: (i) in the general case (arbitrary F), irrespective of whether mutations affect individual or multiple traits, one cannot necessarily expect zero epistasis relative to fitness, even if no epistasis is assumed relative to each individual trait; (ii) when F can be

Table 3. The general expression of epistasis with and without pleiotropy. Equation (3.2) can be rewritten as $\varepsilon = \varepsilon_X + \varepsilon_Y + \varepsilon_{XY}$ (first row), where $\varepsilon_X = \Delta x_i \Delta x_j (\ddot{F}_X/x_0 + \ddot{F}_{XX} - \ddot{F}_X^2)$, $\varepsilon_Y = \Delta y_i \Delta y_j (\ddot{F}_Y/y_0 + \ddot{F}_{YY} - \ddot{F}_Y^2)$ and $\varepsilon_{XY} = (\Delta x_i \Delta y_j + \Delta x_j \Delta y_i) \cdot (\ddot{F}_{XY} - \ddot{F}_X \ddot{F}_Y)$. If each of the alleles i and j acts on a distinct trait with no pleiotropic effect (figure 1*b*; $\Delta x_j = \Delta y_i = 0$, or, equivalently, $\Delta x_i = \Delta y_j = 0$), then one obtains $\varepsilon_X = \varepsilon_Y = 0$, and hence $\varepsilon = \varepsilon_{XY}$. However, for any decomposable function $F(X, Y) = G(X) \cdot H(Y)$ (second row), $\varepsilon_{XY} = 0$ because $\ddot{F}_{XY} - \ddot{F}_X \ddot{F}_Y = 0$. Therefore, when $F(X, Y) = G(X) \cdot H(Y)$, epistasis is non-zero only in the presence of pleiotropy, i.e. if ε_X and/or ε_Y are different from zero. For the particular case $F(X, Y) = X^n Y^m$ (third row), epistasis is always zero, no matter whether or not there is pleiotropy.

	pleiotropic case (figure 1 <i>a</i>)	non-pleiotropic case (figure 1 <i>b</i>)
general $F(X, Y)$	$\varepsilon = \varepsilon_X + \varepsilon_Y + \varepsilon_{XY}$	$\varepsilon = \varepsilon_{XY}$
$F(X, Y) = G(X) \cdot H(Y)$	$\varepsilon = \varepsilon_X + \varepsilon_Y$	$\varepsilon = 0$
$F(X, Y) = X^n Y^m$	$\varepsilon = 0$	$\varepsilon = 0$

decomposed as in equation (3.4), epistasis can exist only in the presence of pleiotropy, i.e. if each mutation affects more than one trait. Hence, equation (3.4) can be viewed as an operational definition of independence between two traits; and (iii) in particular cases (e.g. if F is simply the product of two traits), epistasis will be always zero, irrespective of pleiotropy (table 3).

4. DISCUSSION

Genome-wide epistatic profiles of fitness have been used to study the functional organization of biological systems [10], suggesting the existence of functionally coherent modules, characterized by specific epistatic interaction network properties such as monochromaticity [9,40]. Yet the relationship between the modular organization of the cell and epistasis remains poorly understood. Inspired by the successful attempt to explain epistasis data with a simple benefit–cost model of microbial fitness [7], we sought to explore the broader implications of expressing fitness as a function of two quantitative traits. We found that epistasis, and specific biases in its distribution, could be a natural outcome of the dependence of fitness on multiple phenotypes. If we interpret different phenotypes as metrics associated with different subsystems, or modules, we can determine whether epistasis exists at the system-level, and what might be the bias in its distribution based on how two modules interact with each other to produce fitness. Thus, we suggest that our approach establishes a novel link between biological modularity, pleiotropy and epistasis.

The benefit–cost model, originally employed to explain a small number of interactions in a single evolutionary experiment, has been shown here to lead to a global bias in the distribution of epistasis, under a broad range of beneficial versus deleterious mutation frequencies. In particular, upon deviating from symmetry in the amount of beneficial and deleterious mutations, we predict an overall pattern of negative ε . A similar pattern was previously suggested to be informative in identifying physically interacting partners or gene pairs belonging to redundant, but parallel functional pathways [10]. Besides its relevance to functional genomics, a trend towards negative ε also plays an important role in evolutionary theories seeking to interpret origin and maintenance of sex and recombination [3,4]. On the other hand, for adaptation, negative ε indicates that beneficial alleles combine antagonistically and suggests a diminishing returns trend in the fitness improvements as

more beneficial alleles are acquired, consistent with some previous experimental results and theoretical analyses [7,8,41–43]. Thus, based on simple assumptions about the functional dependence of fitness on multiple phenotypes and genetic perturbations, a single ‘symmetry breaking’ mechanism could provide a potential explanation for both the deceleration of adaptation upon accumulation of beneficial mutations, and the prevalence of synergistic interactions between deleterious mutations. In pondering the general relevance of the above results to biology, one cannot avoid asking whether and why a simple benefit–cost model should truly underlie trends of epistasis, potentially across different organisms and biological scales. Indeed, we do not expect that a benefit–cost model should be the common mechanism behind all observed epistasis trends (see counter-example in Chou *et al.* [44]). However, it is not inconceivable that under some circumstances, fitness could be effectively represented as the overall difference of two independently measurable terms, for example, in cases where most mutations are estimated to affect the abundance of different proteins in the cell. While direct experimental testing is beyond the scope of the current paper, it is worth mentioning that the quantitative prediction of negative epistasis bias in our benefit–cost model does not deviate much from the corresponding value observed for epistatic interactions between yeast deletion mutants [10] upon matching our model’s deleterious/beneficial allele ratio to the 3 : 1 ratio present in the yeast data (see the electronic supplementary material, figure S8). Given that our model is only remotely related to the yeast deletion data, this result may be pure coincidence. However, it exemplifies how the result we are presenting could be tested against experimental data in the future.

Whether or not the negative bias we observe in our model is directly relevant for the discussion on the evolution of recombination remains to be seen. Experimental data have provided conflicting results [23,29,32,45], depending on the system used (yeast, bacteria, viruses), the methods involved (classical versus high throughput), the different criteria for selecting genes to be analysed (highly deleterious only, YPD essential genes, etc.), and the definition of epistasis employed. In addition, while robust with respect to several parameters, the bias observed in the benefit–cost model may still in principle change in magnitude or sign, under different assumptions on the underlying distribution of individual mutations (see also electronic supplementary material, figures S2–S6). In any case, similar to prior computational models [9], we

believe that the framework we are proposing will be useful in explaining, and potentially motivating experimental measurements relevant for this question.

Beyond the benefit–cost model, we showed in equation (3.2) that it is possible to estimate how epistasis depends on the magnitude of individual mutations and on the functional dependence of fitness on the two traits X and Y . For simple functional dependencies, this equation leads to direct insight into the type of epistasis to be expected. For example, linear combinations of traits lead to epistatic effects formally similar to the ones obtained with the benefit–cost model, except that the sign (and therefore the distribution biases) of epistasis is heavily influenced by the signs of the coefficients of such linear combinations. Future experiments (e.g. measuring epistasis upon perturbations of metabolic pathways that combine additively to produce fitness) could directly test this prediction. Our analytical expression is derived upon analysing the partial derivative of fitness with respect to traits it depends on, similar to the concept of ‘phenotype landscape’ proposed before to address the evolution of canalization, phenotypic plasticity and integration [46]. In our approach, however, we explicitly take into account both the effects of different mutations on distinct traits, and the functional dependence of fitness on such traits. One of the most interesting consequences of our general expression for epistasis as a function of two quantitative traits is the possibility to infer a general class of functional dependencies that guarantee lack of epistasis in the absence of pleiotropy. This finding echoes the viewpoint of a recent review that pleiotropy is an important prerequisite for epistasis [24]. Our result establishes a formal link between epistasis and pleiotropy, and suggests a new way to think about independence in biological networks.

Our method is simple and analytically solvable. Future variants of our framework could address more complex or alternative scenarios: first, given the multiplicity of traits that may be thought of as contributing to fitness, one could extend the current approach to fitness functions that depend upon more than two phenotypes (as we preliminarily explored here through computer simulations for the benefit–cost model). Second, while we have assumed here that perturbations correspond to genetic mutations, one could explore the consequences of a similar model for environmental perturbations. Third, it may be interesting to generalize our expression for epistasis to the case in which the basic traits themselves do have some degree of epistasis. In such case, one could seek an ‘epistasis propagation law’, showing how epistasis at a low-level phenotype affects epistasis at higher levels. Fourth, it has been suggested that the purging of deleterious mutations depends on the magnitude of mutational effects, in addition to the bias in the distribution of epistasis [23,47]. Future extension of our equation (3.2) in regard to linkage disequilibrium could further address this view point from another perspective to further elucidate this critical point about the evolution of sexual reproduction. Fifth, as we show in electronic supplementary material, figure S7, our model can account for instances of reciprocal sign epistasis between two alleles, providing potential new avenues for studying the ruggedness of evolutionary landscapes.

Finally, while in the current work we have focused on fitness and on its dependence on other traits, our formulation is quite general, and should equally apply to the functional dependence of any trait on any other set of traits. Even if most genetic interaction data are obtained relative to fitness, epistasis relative to multiple phenotypes will probably become increasingly available [31,48], offering opportunities to study how epistasis propagates between different traits to ultimately shape the genotype–phenotype mapping. We hope also that our equation will help understand epistasis relative to traits associated with genetic diseases, and provide insight on the interplay between evolution and modular organization of biological systems.

We are grateful to Evan Snitkin, Niels Klitgord, Ed Reznik, Nigel Delaney and Deepa Agashe and other members of the Segrè and Marx Laboratories for helpful suggestions and feedback on the manuscript, and to Bjorn Poonen for suggesting a criterion for functional decomposability. This work was supported by NIH grants (R01 GM078209 and R01 GM089978), and by the NASA Astrobiology Institute.

REFERENCES

- Weinreich, D. M., Delaney, N. F., DePristo, M. A. & Hartl, D. L. 2006 Darwinian evolution can follow only very few mutational paths to fitter proteins. *Science* **312**, 111–114. (doi:10.1126/science.1123539)
- Hayden, E. J., Ferrada, E. & Wagner, A. 2011 Cryptic genetic variation promotes rapid evolutionary adaptation in an RNA enzyme. *Nature* **474**, 92–95. (doi:10.1038/nature10083)
- Kondrashov, A. S. 1988 Deleterious mutations and the evolution of sexual reproduction. *Nature* **336**, 435–440. (doi:10.1038/336435a0)
- Otto, S. P. 2009 The evolutionary enigma of sex. *Am. Nat.* **174**, S1–S14. (doi:10.1086/599084)
- Sanjuan, R., Cuevas, J. M., Moya, A. & Elena, S. F. 2005 Epistasis and the adaptability of an RNA virus. *Proc. Natl Acad. Sci. USA* **170**, 1001–1008. (doi:10.1534/genetics.105.040741)
- Kryazhimskiy, S., Tkačik, G. & Plotkin, J. B. 2009 The dynamics of adaptation on correlated fitness landscapes. *Proc. Natl. Acad. Sci. USA* **106**, 18 638–18 643. (doi:10.1073/pnas.0905497106)
- Chou, H.-H., Chiu, H.-C., Delaney, N. F., Segrè, D. & Marx, C. J. 2011 Diminishing returns epistasis among beneficial mutations decelerates adaptation. *Science* **332**, 1190–1192. (doi:10.1126/science.1203799)
- Kahn, A. I., Dinh, D. M., Schneider, D., Lenski, R. E. & Cooper, T. F. 2011 Negative epistasis between beneficial mutations slows the rate of population adaptation. *Science* **332**, 1193–1196. (doi:10.1126/science.1203801)
- Segrè, D., DeLuna, A., Church, G. M. & Kishony, R. 2005 Modular epistasis in yeast metabolism. *Nat. Genet.* **37**, 77–83. (doi:10.1038/ng1489)
- Costanzo, M. *et al.* 2010 The genetic landscape of a cell. *Science* **327**, 425–431. (doi:10.1126/science.1180823)
- Hartwell, L. H., Hopfield, J. J., Leibler, S. & Murray, A. W. 1999 From molecular to modular cell biology. *Nature* **402**, C47–C52. (doi:10.1038/35011540)
- Lawrence, J. G. & Roth, J. R. 1996 Selfish operons: horizontal transfer may drive the evolution of gene clusters. *Genetics* **143**, 1843–1860.
- Kreimer, A., Borenstein, E., Gophna, U. & Ruppin, E. 2008 The evolution of modularity in bacterial metabolic networks. *Proc. Natl Acad. Sci. USA* **105**, 6976–6981. (doi:10.1073/pnas.0712149105)

- 14 Dekel, E. & Alon, U. 2005 Optimality and evolutionary tuning of the expression level of a protein. *Nature* **436**, 588–592. (doi:10.1038/nature03842)
- 15 Kashtan, N. & Alon, U. 2005 Spontaneous evolution of modularity and network motifs. *Proc. Natl Acad. Sci. USA* **102**, 13 773–13 778. (doi:10.1073/pnas.0503610102)
- 16 Wagner, G. P., Pavlicev, M. & Cheverud, J. M. 2007 The road to modularity. *Nat. Rev. Genet.* **8**, 921–931. (doi:10.1038/nrg2267)
- 17 Wade, M. J., Winther, R. G., Agrawal, A. F. & Goodnight, C. J. 2001 Alternative definitions of epistasis: dependence and interaction. *Trends Ecol. Evol.* **16**, 498–504. (doi:10.1016/S0169-5347(01)02213-3)
- 18 Cordell, H. J. 2002 Epistasis: what it means, what it doesn't mean, and statistical methods to detect it in humans. *Hum. Mol. Genet.* **11**, 2463–2468. (doi:10.1093/hmg/11.20.2463)
- 19 Mani, R., St Onge, R. P., Hartman IV, J. L., Giaever, G. & Roth, F. P. 2008 Defining genetic interaction. *Proc. Natl Acad. Sci. USA* **105**, 3461–3466. (doi:10.1073/pnas.0712255105)
- 20 Gao, H., Granka, J. M. & Feldman, M. W. 2010 On the classification of epistatic interactions. *Genetics* **184**, 827–837. (doi:10.1534/genetics.109.111120)
- 21 Loewe, L. & Hill, W. G. 2010 The population genetics of mutations: good, bad and indifferent. *Phil. Trans. R. Soc. B* **365**, 1153–1167. (doi:10.1098/rstb.2009.0317)
- 22 Phillips, P. C., Otto, S. P. & Whitlock, M. C. 2000 Beyond the average: the evolutionary importance of gene interactions and variability of epistatic effects. In *Epistasis and the evolutionary process* (eds J. B. Wolf, E. D. Brodie & M. J. Wade), pp. 20–38. New York, NY: Oxford University Press.
- 23 Kouyos, R. D., Silander, O. K. & Bonhoeffer, S. 2007 Epistasis between deleterious mutations and the evolution of recombination. *Trends Ecol. Evol.* **22**, 308–315. (doi:10.1016/j.tree.2007.02.014)
- 24 de Visser, J. A. G. M., Cooper, T. F. & Elena, S. F. 2011 The causes of epistasis. *Proc. R. Soc. B* **278**, 3617–3624. (doi:10.1098/rspb.2011.1537)
- 25 Elena, S. F. & Lenski, R. E. 1997 Test of synergistic interactions among deleterious mutations in bacteria. *Nature* **390**, 395–398. (doi:10.1038/37108)
- 26 de Visser, J. A. G. M., Hoekstra, R. F. & van den Ende, H. 1997 Test of interaction between genetic markers that affect fitness in *Aspergillus niger*. *Evolution* **51**, 1499–1505. (doi:10.2307/2411202)
- 27 Wloch, D. M., Borts, R. H. & Korona, R. 2001 Epistatic interactions of spontaneous mutations in haploid strains of the yeast *Saccharomyces cerevisiae*. *J. Evol. Biol.* **14**, 310–316. (doi:10.1046/j.1420-9101.2001.00267.x)
- 28 Szafraniec, K., Wloch, D. M., Sliwa, P., Borts, R. H. & Korona, R. 2003 Small fitness effects and weak genetic interactions between deleterious mutations in heterozygous loci of the yeast *Saccharomyces cerevisiae*. *Genet. Res.* **82**, 19–31. (doi:10.1017/S001667230300630X)
- 29 Jasnos, L. & Korona, R. 2007 Epistatic buffering of fitness loss in yeast double deletion strains. *Nat. Genet.* **39**, 550–554. (doi:10.1038/ng1986)
- 30 Martin, G., Elena, S. F. & Lenormand, T. 2007 Distributions of epistasis in microbes fit predictions from a fitness landscape model. *Nat. Genet.* **39**, 555–560. (doi:10.1038/ng1998)
- 31 Snitkin, E. S. & Segrè, D. 2011 Epistatic interaction maps relative to multiple metabolic phenotypes. *PLoS Genet.* **7**, e1001294. (doi:10.1371/journal.pgen.1001294)
- 32 He, X., Qian, W., Wang, Z., Li, Y. & Zhang, J. 2010 Prevalent positive epistasis in *Escherichia coli* and *Saccharomyces cerevisiae* metabolic networks. *Nat. Genet.* **42**, 272–276. (doi:10.1038/ng.524)
- 33 Kryazhimskiy, S., Draghi, J. A. & Plotkin, J. B. 2011 In evolution, the sum is less than its parts. *Science* **332**, 1160–1161. (doi:10.1126/science.1208072)
- 34 Tănase-Nicola, S. & ten Wolde, P. R. 2008 Regulatory control and the costs and benefits of biochemical noise. *PLoS Comput. Biol.* **4**, e1000125. (doi:10.1371/journal.pcbi.1000125)
- 35 Wessely, F., Bartl, M., Guthke, R., Li, P., Schuster, S. & Kaleta, C. 2011 Optimal regulatory strategies for metabolic pathways in *Escherichia coli* depending on protein costs. *Mol. Syst. Biol.* **7**. (doi:10.1038/msb.2011.46)
- 36 Kondrashov, F. A. & Kondrashov, A. S. 2001 Multi-dimensional epistasis and the disadvantage of sex. *Proc. Natl Acad. Sci. USA* **98**, 12 089–12 092. (doi:10.1073/pnas.211214298)
- 37 Pfeiffer, T. & Bonhoeffer, S. 2002 Evolutionary consequences of tradeoffs between yield and rate of ATP production. *Z. Phys. Chem.* **216**, 51. (doi:10.1524/zpch.2002.216.1.051)
- 38 Sanjuán, R., Moya, A. & Elena, S. F. 2004 The contribution of epistasis to the architecture of fitness in an RNA virus. *Proc. Natl Acad. Sci. USA* **101**, 15 376–15 379. (doi:10.1073/pnas.0404125101)
- 39 Phillips, P. C. 2008 Epistasis: the essential role of gene interactions in the structure and evolution of genetic systems. *Nat. Rev. Genet.* **9**, 855–867. (doi:10.1038/nrg2452)
- 40 Michaut, M., Baryshnikova, A., Costanzo, M., Myers, C. L., Andrews, B. J., Boone, C. & Bader, G. D. 2011 Protein complexes are central in the yeast genetic landscape. *PLoS Comput. Biol.* **7**, e1001092. (doi:10.1371/journal.pcbi.1001092)
- 41 Gerrish, P. J. & Lenski, R. E. 1998 The fate of competing beneficial mutations in an asexual population. *Genetica* **102–103**, 127–144. (doi:10.1023/A:1017067816551)
- 42 de Visser, J. A. G. M., Zeyl, C. W., Gerrish, P. J., Blanchard, J. L. & Lenski, R. E. 1999 Diminishing returns from mutation supply rate in asexual populations. *Science* **283**, 404–406. (doi:10.1126/science.283.5400.404)
- 43 Elena, S. F. & Lenski, R. E. 2003 Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. *Nat. Rev. Genet.* **4**, 457–469. (doi:10.1038/nrg1088)
- 44 Chou, H.-H., Berthet, J. & Marx, C. J. 2009 Fast growth increases the selective advantage of a mutation arising recurrently during evolution under metal limitation. *PLoS Genet.* **5**, e1000652. (doi:10.1371/journal.pgen.1000652)
- 45 Bonhoeffer, S., Chappey, C., Parkin, N. T., Whitcomb, J. M. & Petropoulos, C. J. 2004 Evidence for positive epistasis in HIV-1. *Science* **306**, 1547–1550. (doi:10.1126/science.1101786)
- 46 Rice, S. H. 2000 The evolution of developmental interactions: epistasis, canalization, and integration. In *Epistasis and the evolutionary process* (eds J. B. Wolf, E. D. Brodie & M. J. Wade), pp. 82–98. New York, NY: Oxford University Press.
- 47 Kouyos, R. D., Otto, S. P. & Bonhoeffer, S. 2006 Effect of varying epistasis on the evolution of recombination. *Genetics* **173**, 589–597. (doi:10.1534/genetics.105.053108)
- 48 Rowe, H. C., Hansen, B. G., Halkier, B. A. & Kliebenstein, D. J. 2008 Biochemical networks and epistasis shape the *Arabidopsis thaliana* metabolome. *Plant Cell* **20**, 1199–1216. (doi:10.1105/tpc.108.058131)

Supplementary Material for the manuscript

Chiu, Marx and Segrè

Epistasis from functional dependence of fitness on underlying traits

Content

Supplementary Figures	pp. 2-10
Supplementary Methods	pp. 11-20

Supplementary Figures

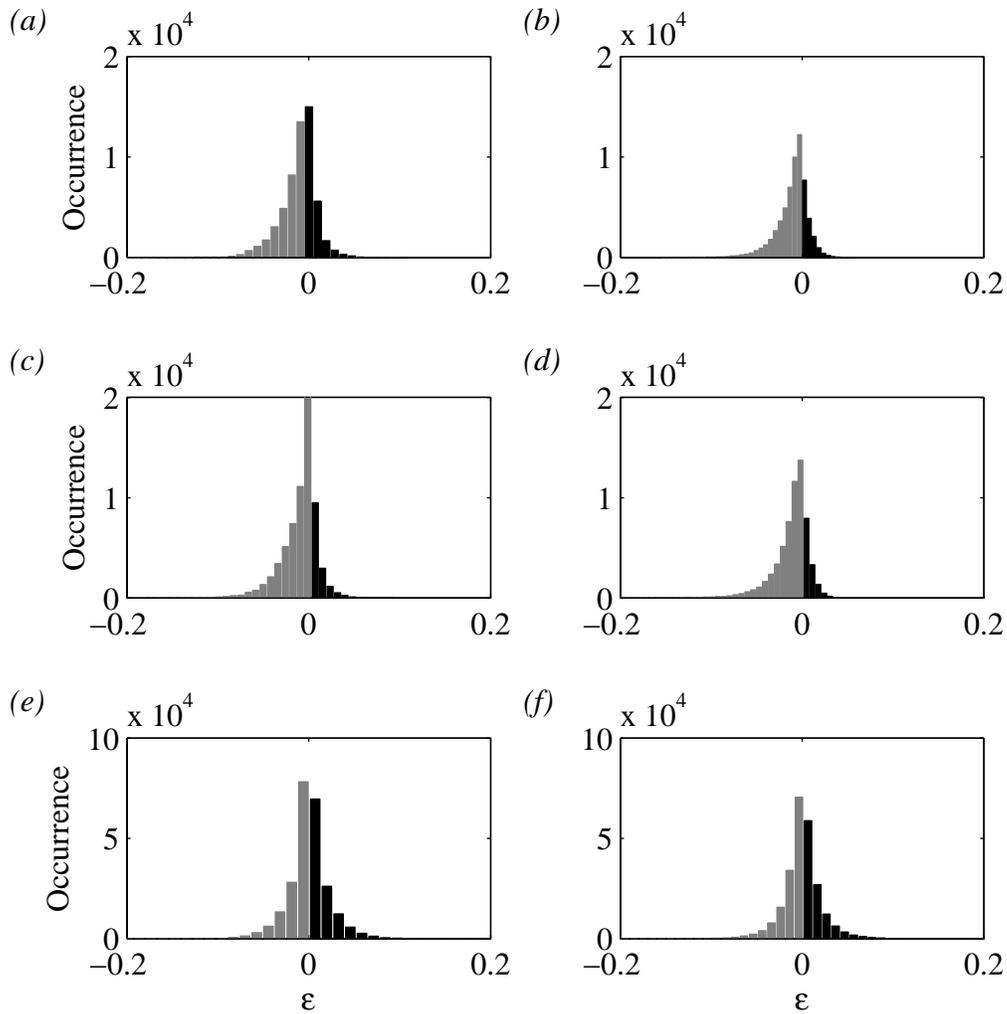


Figure S1. The negative epistasis trend observed in the benefit-cost model (figure 2) is also found in more complex extensions that include multiple cost terms (See Supplementary Methods A). The left panels correspond to a two cost system ($f_0 = b_0 - c_0 - d_0$), while the right panels refer to three cost system ($f_0 = b_0 - c_0 - d_0 - e_0$). (a-b) Strictly beneficial alleles. (c-d) Strictly deleterious alleles (e-f) No constraint on whether alleles are beneficial or deleterious.

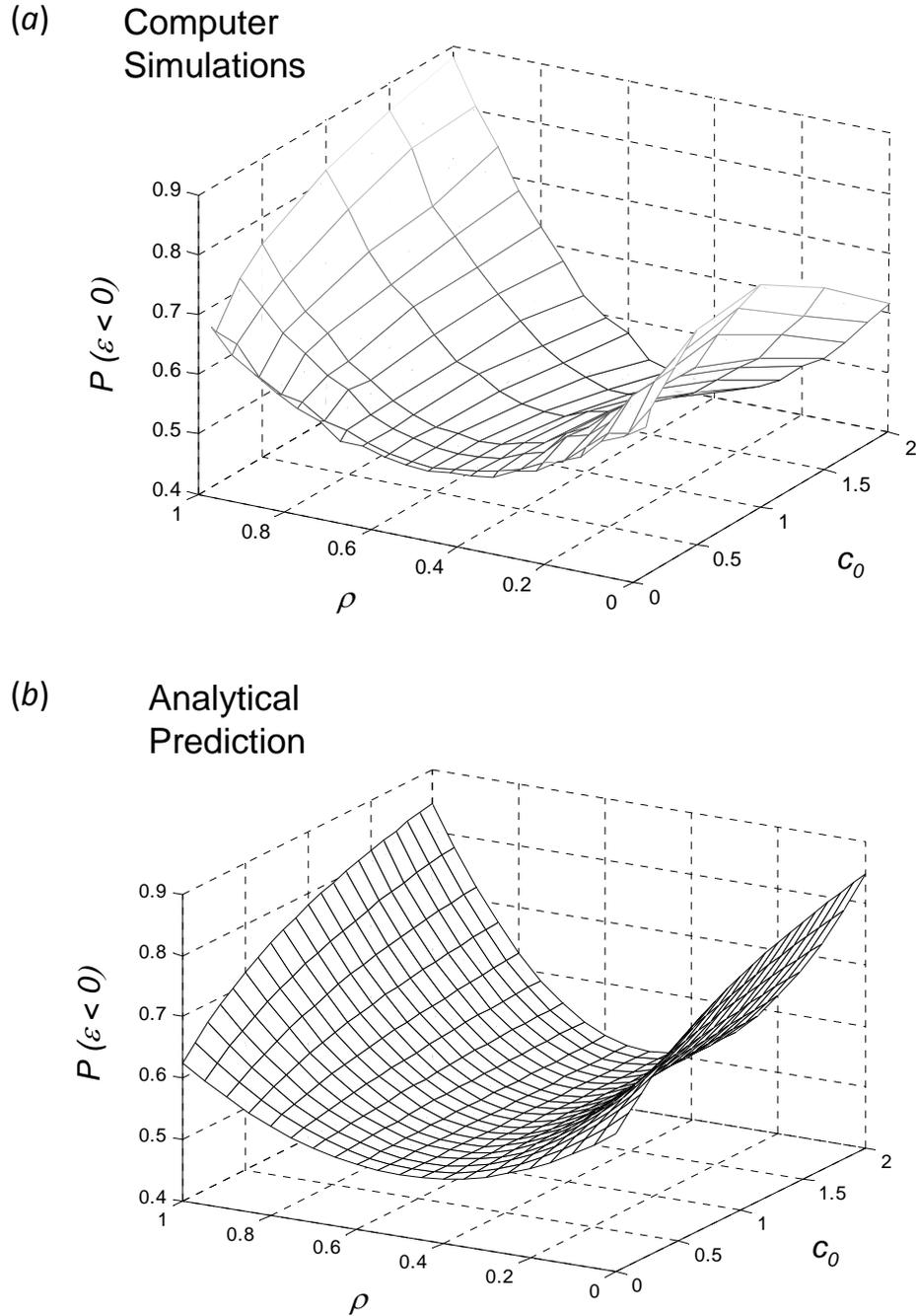


Figure S2. A global sensitivity analysis of the bias in epistasis was performed as a function of the c_0 parameter in the benefit-cost model, and as a function of the imposed chance of beneficial mutations ρ . (a) Surface obtained from computer simulations, as explained in Supplementary Methods A. (b) Surface obtained from analytical calculations (see Supplementary Methods C). Note that here we plot $p(\varepsilon < 0)$, whereas in our analytical calculations we write an expression for Δp^ε . Due to distribution normalization, the two are related by the simple relationship: $\Delta p^\varepsilon = 2p(\varepsilon < 0) - 1$.

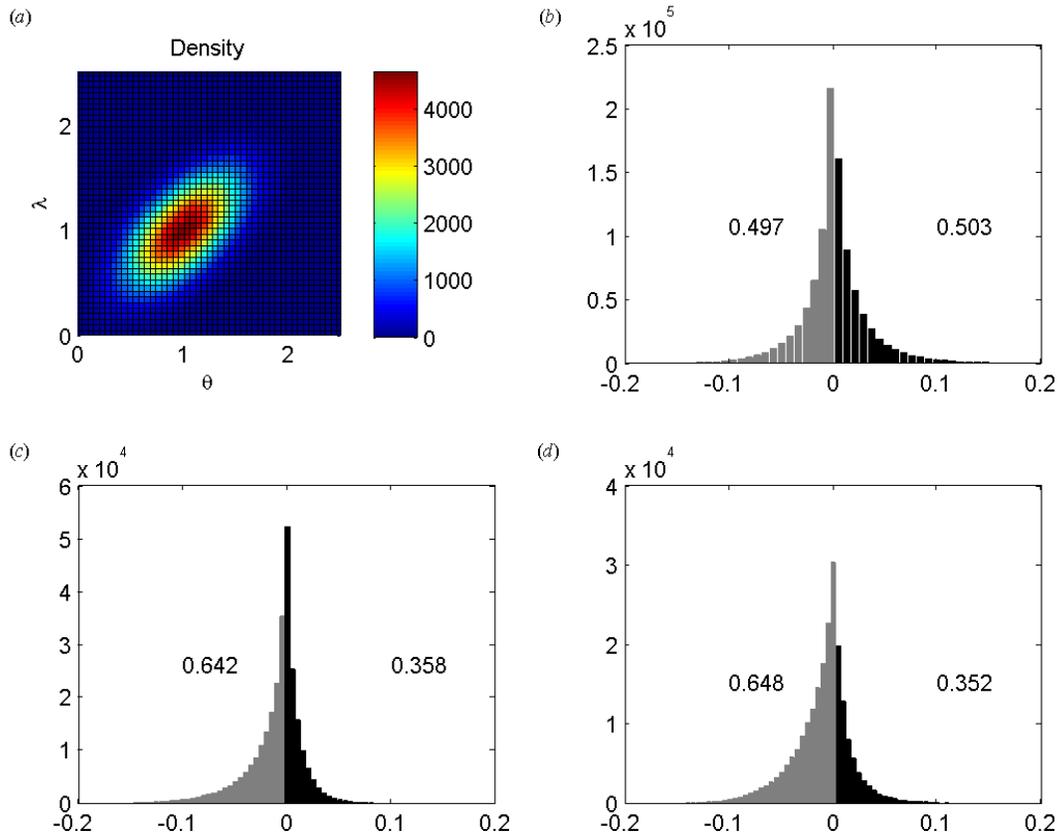


Figure S3. Distributions of epistasis for pairs of mutations whose λ and θ values are inherently correlated. Interdependence between λ and θ is modeled by sampling from a rotated bivariate distribution in the (λ, θ) plane. (a) The bivariate probability density is shown here under a rotation angle of $\pi \cdot 3/4$, with $\sigma_\lambda = 0.05$, $\sigma_\theta = 0.1$. Based on this distribution of λ and θ , we generated single alleles and analyzed epistasis between them. In general, despite the introduced correlations, we observe epistasis biases similar to the ones found for uncorrelated samplings (e.g.: figure 2 of the main text). (b) Without imposing any constraint on whether mutations are beneficial or deleterious an unbiased epistasis distribution is observed. (c-d) Negative epistasis trends are observed between two strictly beneficial alleles (panel c), and between two strictly deleterious alleles (panel d). Analyses with other rotation angles (see figures S4-6) also suggest similar results.

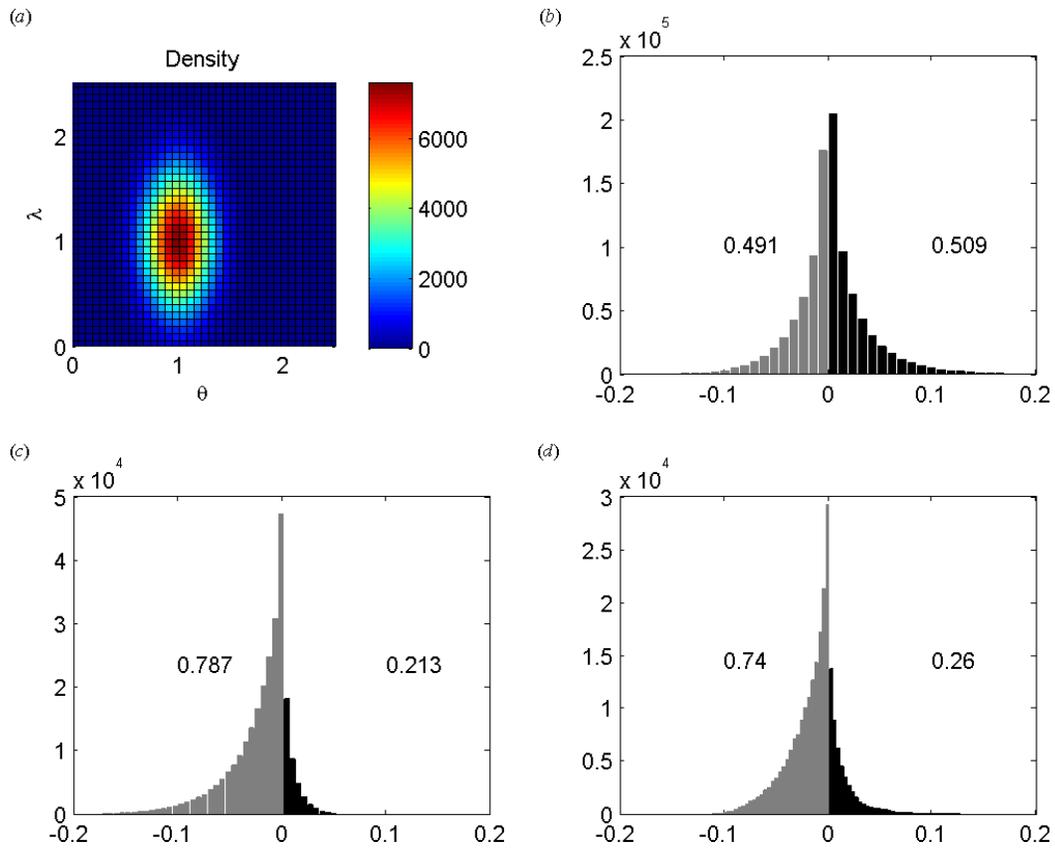


Figure S4. Similar analysis to figure S3 with rotation angle = 0. See legend of figure S3 for details.

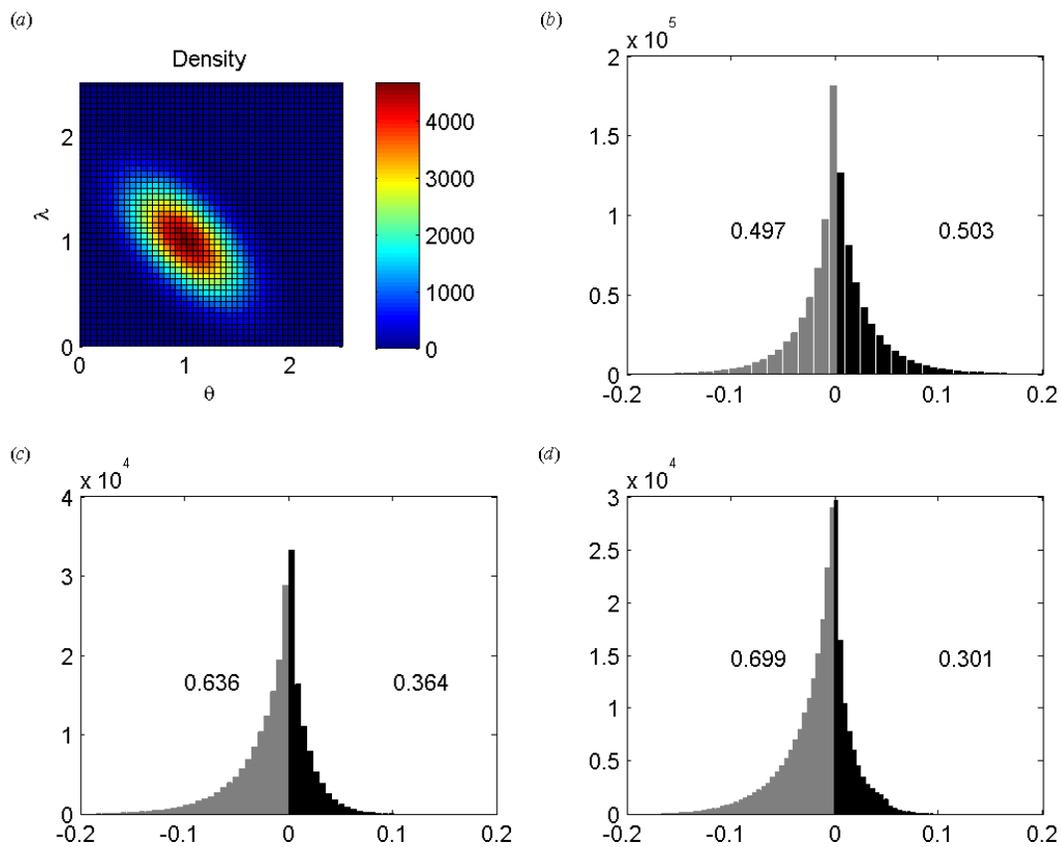


Figure S5. Similar analysis to figure S3 with rotation angle = $\pi/4$. See legend of figure S3 for details.

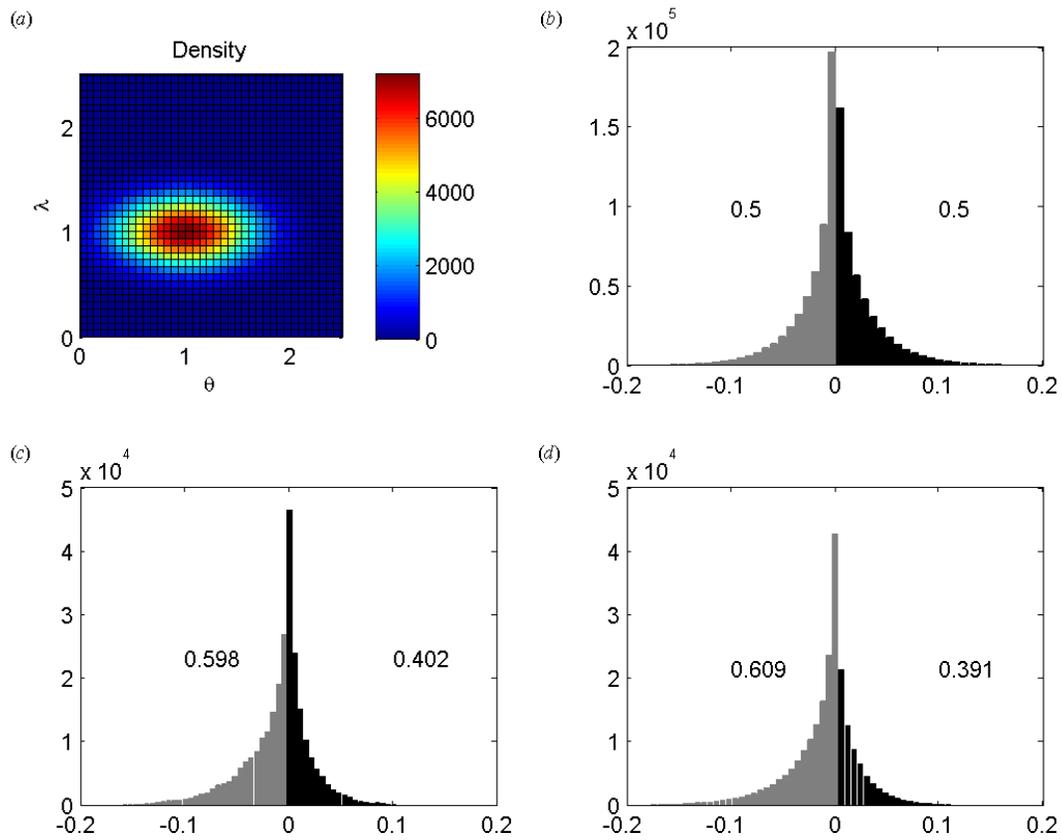


Figure S6. Similar analysis to figure S3 with rotation angle = $\pi/2$. See legend of figure S3 for details.

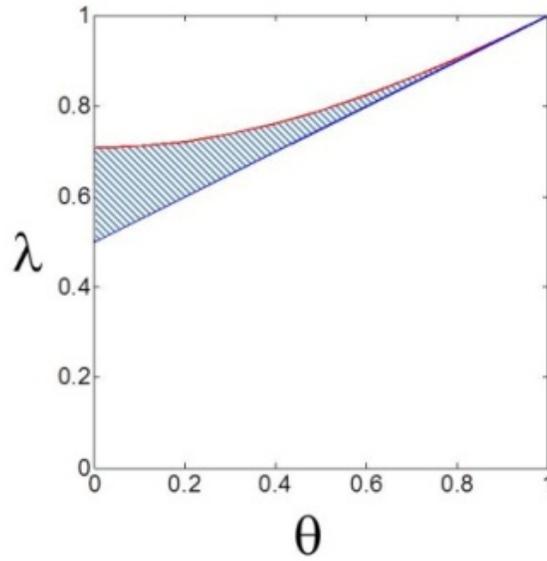


Figure S7. A region in the λ, θ parameter space that gives rise to sign epistasis in the benefit-cost model. While it may not be trivial to characterize all possible sets of parameters that would give rise to sign epistasis, it is possible to identify a subset of such parameter sets through a simple analytical derivation. If we choose a subset of possible mutations such that $\lambda_i = \lambda_j = \lambda$ and $\theta_i = \theta_j = \theta$, a condition for observing reciprocal sign epistasis is that $\lambda b_0 - \theta c_0 > 1$ while $\lambda^2 b_0 - \theta^2 c_0 < 1$. The region satisfying these two conditions is shown in the blue shaded area of the figure. An example of perturbations that would give rise to sign epistasis is the case of $b_0 = 2$ and $c_0 = 1$, with both $\lambda_i = \lambda_j = 0.6$ and $\theta_i = \theta_j = 0.1$. In this case, the fitness of single mutants would be $f_i = f_j = 0.6 * 2 - 0.1 * 1 = 1.1$, such that both are individually beneficial. Combined, however, they are net deleterious: $f_{ij} = 0.6 * 0.6 * 2 - 0.1 * 0.1 * 1 = 0.71$.

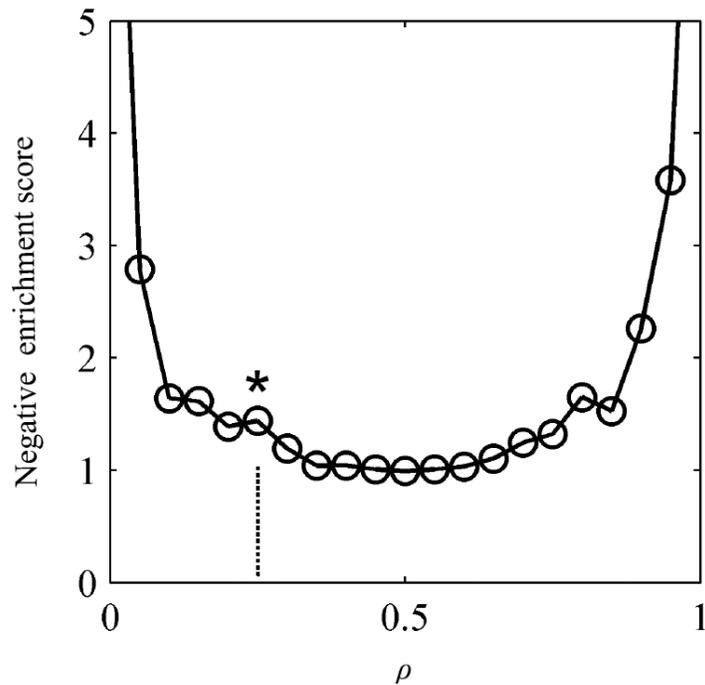


Figure S8. Negative enrichment score (total probability of significantly negative ε divided by total probability of significantly positive ε) indicates that our model (1.4 fold when beneficial percentage is 0.25, dashed line) qualitatively agrees with the experimental results (approximately 2 fold, asterisk) observed in the yeast deletion data [10]. A different type of comparison yields a 1% difference between our model and the yeast data [10] for the fraction of negative ε (Supplementary Methods A).

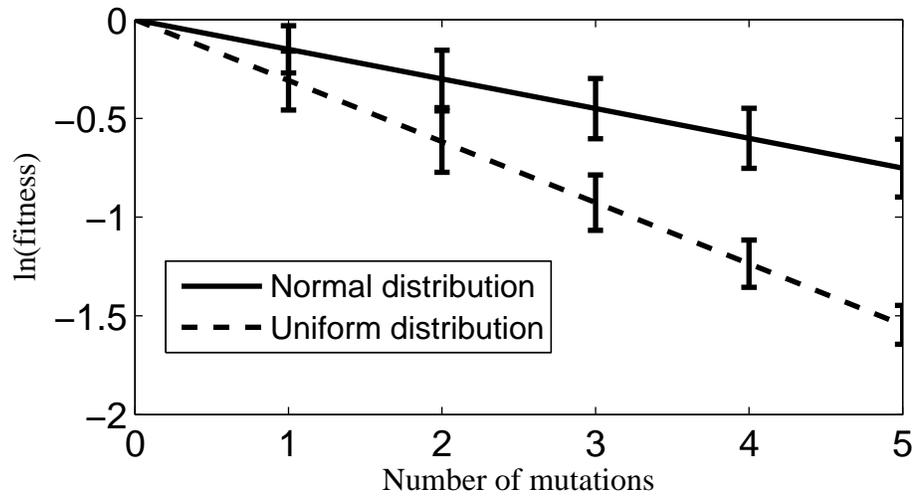


Figure S9. Predicted effect on fitness of the accumulation of multiple deleterious mutations that display lack of epistasis with each other (due to identical choice of mutational effects on benefit and cost, i.e. $\lambda_i = \theta_i$ for any mutation i). Under these assumptions, we verified that fitness decreases log-linearly with the number of mutations, as expected (see Supplementary Methods C for additional explanations).

Supplementary Methods

A. Numerical estimate of epistasis distributions under the benefit-cost model

In the benefit-cost model, each mutant allele corresponds to two factors (λ_i and θ_i) that multiplicatively modify the benefit (b_0) and cost (c_0) respectively. Equation (5) from the main text allows us to evaluate epistasis between any pair of mutations in this model. In addition to using this model for analytical estimates of biases in the distribution of epistasis, we have implemented numerical simulations that allow us to plot the full distributions, and estimate dependence on parameters. In this method section we describe how we perform the sampling of alleles and calculation of epistasis, used in figure 2*b-d*, figure 3, figure S1, figure S2*a*, and figure S8.

General model settings

For continuity with our previous work [7], we use a benefit-cost model with experimentally measured parameters $b_0 = 1.141$, $c_0 = 0.141$. Single mutant alleles are chosen by sampling both λ_i and θ_i from $Normal(1, 0.3)$. A natural lower bound of zero is enforced for both λ and θ because negative values for such parameters are biologically meaningless. In order to maintain symmetric distributions of these parameters an upper bound of $W=2$ is also imposed to their values (see figure 2). These parameters are used to generate a large set of randomly sampled single alleles. By building all possible pairs of single mutants, we can compute fitness for double mutants (as defined by equation 4), and epistasis (as in equation 1). All double mutants are used to draw the epistasis distribution in figure 2*b*. Two subsets, one of strictly beneficial alleles (figure 2*c*) and one of strictly deleterious alleles (figure 2*d*) are selected among all double mutants to draw corresponding figure panels.

Numerical computation of full epistasis distributions

The aim of the simulation presented in figure 3 is to produce the distribution of epistasis as a function of the fraction of beneficial alleles in the population (labeled as ' ρ ' in figure 3*a*). Thus, for each value of ρ (the fraction of beneficial alleles, where ρ varies between 0 and 1 in intervals of 0.05), we generate (as described above) a set of single mutants of size 2500 satisfying the condition that the fraction of beneficial allele is equal to ρ . For instance, if $\rho = 0.4$, we generate 1000 beneficial and 1500 deleterious single mutants. We then generate all possible double mutants, calculate ε and plot the distribution of epistasis as a function of ρ . For each value of ρ we generate a "slice" of the multidimensional epistasis distribution shown in figure 3*a* (grey for negative ε and dark for positive ε) with different degree of bias. The data size (2500×2500) is chosen so as to provide results comparable with the large scale epistasis distribution that is experimentally measured in [10].

In order to estimate the dependence of the bias in epistasis as a function of the c_0 parameter, in addition to ρ , we repeated the above procedure for different values of c_0 . This global sensitivity analysis allowed us to compute a surface (figure S2*a*) that can be directly compared to the results of analytical derivations (figure S2*b* and Supplementary Methods C).

Negative epistasis bias in multi-cost fitness system

To test whether the overall negative bias is sensitive to the number N_c of cost components in the benefit-cost model we implemented multi-cost systems. In these variants of the basic

model, each mutant allele is represented by a λ factor and $N_c=2$ or $N_c=3$ θ factors. Random sampling in this case is a straightforward extension of the procedure described above. For two-cost systems, ($f_0 = b_0 - c_0 - d_0$, figure S1 first column) and three-cost systems ($f_0 = b_0 - c_0 - d_0 - e_0$, figure S1 second column), negative bias remains with a general property of ε , as in the results of figure 2. For each simulation, the total cost (e.g. $c_0 + d_0$) is controlled to be the same as the one used in the original benefit-cost system [7] so that the number of the cost components is the only variable tested.

Comparing benefit-cost predictions of epistasis with yeast deletion experimental results

We compared our theoretical prediction of negative bias derived from the benefit-cost model with the experimentally inferred bias in epistatic interactions between yeast deletion mutants [10]. By revisiting the raw data provided in a large compendium of yeast single and double gene deletion growth phenotypes [10], we estimated the fraction of negative ε . In particular, we computed the number of interactions with $\varepsilon < 0$ divided by total interactions having valid ε from their supplementary file S1, and found that the experimentally determined fraction of negative ε , in this case, is 53.4%. We computed the corresponding prediction of the benefit-cost model by imposing a beneficial proportion of alleles $\rho=0.25$ (the same observed in the yeast data). The ensuing fraction of negative ε predicted by the benefit-cost model (54.3%, figure 3b) is surprisingly close to the experimentally derived value, despite the big difference between the two approaches. We also calculated in the benefit-cost model the negative enrichment score (total probability of significantly negative ε divided by total probability of significantly positive ε as reported in [10]). A simple mean ± 3 standard deviation threshold is chosen for determining these significantly negative and positive ε in our calculation. As shown in figure S8, we obtained a 1.4 fold enrichment score, to be compared with the 2 fold enrichment reported in [10].

B. Derivation for epistasis score ε in the benefit-cost model

In the benefit-cost fitness model, $f_0 = b_0 - c_0 = 1$, $f_i = \lambda_i b_0 - \theta_i c_0$, $f_j = \lambda_j b_0 - \theta_j c_0$ and $f_{ij} = \lambda_i \lambda_j b_0 - \theta_i \theta_j c_0$. Thus, the degree of epistasis ε can be derived as:

$$\begin{aligned}
\varepsilon &= \frac{f_{ij}}{f_0} - \frac{f_i}{f_0} \cdot \frac{f_j}{f_0} \\
&= \frac{(\lambda_i \lambda_j b_0 - \theta_i \theta_j c_0)}{b_0 - c_0} - \frac{(\lambda_i b_0 - \theta_i c_0) \cdot (\lambda_j b_0 - \theta_j c_0)}{(b_0 - c_0)^2} \\
&= \frac{(\lambda_i \lambda_j b_0 - \theta_i \theta_j c_0)(b_0 - c_0)}{(b_0 - c_0)^2} - \frac{(\lambda_i \lambda_j b_0^2 - \lambda_i \theta_j b_0 c_0 - \lambda_j \theta_i b_0 c_0 + \theta_i \theta_j c_0^2)}{(b_0 - c_0)^2} \\
&= \frac{1}{(b_0 - c_0)^2} [(\lambda_i \lambda_j b_0^2 - \lambda_i \lambda_j b_0 c_0 - \theta_i \theta_j b_0 c_0 + \theta_i \theta_j c_0^2) - (\lambda_i \lambda_j b_0^2 - \lambda_i \theta_j b_0 c_0 - \lambda_j \theta_i b_0 c_0 + \theta_i \theta_j c_0^2)] \\
&= \frac{1}{(b_0 - c_0)^2} [-\lambda_i \lambda_j b_0 c_0 - \theta_i \theta_j b_0 c_0 + \lambda_i \theta_j b_0 c_0 + \lambda_j \theta_i b_0 c_0] \\
&= \frac{1}{(b_0 - c_0)^2} [b_0 c_0 (-\lambda_i \lambda_j - \theta_i \theta_j + \lambda_i \theta_j + \lambda_j \theta_i)] \\
&= \frac{b_0 c_0}{(b_0 - c_0)^2} [\lambda_i (\theta_j - \lambda_j) + \theta_i (\lambda_j - \theta_j)] \\
&= \frac{b_0 c_0}{(b_0 - c_0)^2} [-(\lambda_i - \theta_i)(\lambda_j - \theta_j)] \\
&= -b_0 c_0 [(\lambda_i - \theta_i)(\lambda_j - \theta_j)]
\end{aligned}$$

As shown above, the sign of ε is determined by the sign of $(\lambda - \theta)$ for allele i and j . In statistical terms, therefore, whether ε is more likely to be positive or negative is determined by the distribution of δ . If we treat allele i and j as independent events coming from the same population where $p(\delta_i > 0) = p(\delta_j > 0) = a$, $p(\delta_i < 0) = p(\delta_j < 0) = b$, we can classify ε into four categories according to the sign of δ_i and δ_j . These categories are analogous to the head and tail events for tossing a coin twice, known as Bernoulli trials, if we treat the sign of δ as a binomial random process. As shown in Table 2, we obtain the prediction that there is a tendency of negative ε as long as δ is biased. In other words:

$$a \neq b \Rightarrow p(\varepsilon < 0) > p(\varepsilon > 0)$$

When δ is unbiased ($a = b$, i.e. δ is equally likely to be positive or negative), negative and positive ε are equally likely to be observed.

C. Negative bias in the distribution of ε for the benefit-cost model

Clarifications about the (λ, θ) plane of figure 2a

Our (λ, θ) plane representation of fitness and mutations is meant to help predict the sign of epistasis between any two arbitrary mutations based on the benefit cost model. While in evolutionary genetics epistasis is often illustrated by plotting fitness as a function of the number of mutations, we wish to emphasize that the (λ, θ) plane has a completely different meaning. The (λ, θ) representation is not meant to provide a map of possible epistasis trends (as one would represent in classical fitness vs. number of mutations graphs). Rather, it serves as a geometrical aid for obtaining a prediction of how likely it is to choose a pair of mutations that will result in epistasis of a given sign. The (λ, θ) plane (and, similarly, Eq. 5) does not lend itself in a straightforward way to extrapolating what would happen in the case of accumulation of an arbitrary number of mutations.

Despite all this, it is possible to ask how the benefit-cost model behaves upon the accumulation of multiple mutations. In particular, it is important to verify what should be a natural trend of our model in absence of epistasis. As reported in a number of other studies [25, 26, 47], if alleles combine non-epistatically, accumulating deleterious alleles should yield a log linear decay in mean fitness as a function of the number of mutations. To verify this trend in our model, we randomly generated 100 deleterious single mutants in which any two single mutants combine non-epistatically (i.e. for any single allele i , $\lambda_i = \theta_i$). Next, we generate all possible combinations of double, triple, quadruple, quintuple mutants and plot the logarithm of average fitness in each case, as a function of the number of mutations. The results are shown in figure S9, under two different assumptions about the distribution of mutational effects. In both cases one can observe the expected log linear decay in average fitness as a function of the number of mutations.

Expected epistasis bias from computational simulation and analytical derivation

In this section, we are going to compute analytically the expected probability of obtaining positive vs. negative ε for possible combinations of mutations chosen in the (λ, θ) plane, for a given chance of beneficial mutations, ρ . Within each of the two regions (beneficial, deleterious) defined by the neutrality line, we assume a uniform probability of choosing λ, θ points.

Let us first focus on the case of two beneficial mutations. The chance of picking two beneficial mutations is ρ^2 . The sign of ε , as discussed for equation (5) in the main text, will depend on whether the chosen beneficial mutations lie in the region above or under the isochange line (figure 2a). In particular, we could have one of four possible cases:

Allele 1 (Benef.)	Allele 2 (Benef.)	Probability of event	Sign(ε)
above	above	$\rho^2 \cdot [B_a / (B_a + B_w)] \cdot [B_a / (B_a + B_w)] = \rho^2 \cdot B_a^2 / (B_a + B_w)^2$	-
above	under	$\rho^2 \cdot [B_a / (B_a + B_w)] \cdot [B_w / (B_a + B_w)] = \rho^2 \cdot B_a B_w / (B_a + B_w)^2$	+
under	above	$\rho^2 \cdot [B_w / (B_a + B_w)] \cdot [B_a / (B_a + B_w)] = \rho^2 \cdot B_w B_a / (B_a + B_w)^2$	+
under	under	$\rho^2 \cdot [B_w / (B_a + B_w)] \cdot [B_w / (B_a + B_w)] = \rho^2 \cdot B_w^2 / (B_a + B_w)^2$	-

Hence, for a combination of beneficial mutations, one can compute:

$$\begin{aligned}
\Delta p^\varepsilon(B,B) &= p(\varepsilon < 0) - p(\varepsilon > 0) = & (SC1) \\
&= [\rho^2/(B_a+B_u)^2] \cdot [B_a^2 + B_u^2 - 2 B_u B_a] = \\
&= \rho^2(B_a-B_u)^2/(B_a+B_u)^2
\end{aligned}$$

Similarly, we can compute a table of all possible outcomes of deleterious mutations (i.e. points under the neutrality line):

Allele 1 (Delet.)	Allele 2 (Delet.)	Probability of event	Sign(ε)
above	above	$(1-\rho)^2 \cdot [D_a/(D_a+D_u)] \cdot [D_a/(D_a+D_u)] = (1-\rho)^2 \cdot D_a^2/(D_a+D_u)^2$	-
above	under	$(1-\rho)^2 \cdot [D_a/(D_a+D_u)] \cdot [D_u/(D_a+D_u)] = (1-\rho)^2 \cdot D_a D_u/(D_a+D_u)^2$	+
under	above	$(1-\rho)^2 \cdot [D_u/(D_a+D_u)] \cdot [D_a/(D_a+D_u)] = (1-\rho)^2 \cdot D_u D_a/(D_a+D_u)^2$	+
under	under	$(1-\rho)^2 \cdot [D_u/(D_a+D_u)] \cdot [D_u/(D_a+D_u)] = (1-\rho)^2 \cdot D_u^2/(D_a+D_u)^2$	-

The ensuing excess chance of negative ε for a combination of deleterious mutations is then:

$$\begin{aligned}
\Delta p^\varepsilon(D,D) &= [(1-\rho)^2/(D_a+D_u)^2] \cdot [D_a^2 + D_u^2 - 2 D_u D_a] = & (SC2) \\
&= (1-\rho)^2 \cdot (D_a-D_u)^2/(D_a+D_u)^2
\end{aligned}$$

Finally, we have to consider the case in which one mutation is beneficial, and the other deleterious:

Allele 1 (Benef.)	Allele 2 (Delet.)	Probability of event	Sign(ε)
above	above	$\rho(1-\rho) \cdot [B_a/(B_a+B_u)] \cdot [D_a/(D_a+D_u)]$	-
above	under	$\rho(1-\rho) \cdot [B_a/(B_a+B_u)] \cdot [D_u/(D_a+D_u)]$	+
under	above	$\rho(1-\rho) \cdot [B_u/(B_a+B_u)] \cdot [D_a/(D_a+D_u)]$	+
under	under	$\rho(1-\rho) \cdot [B_u/(B_a+B_u)] \cdot [D_u/(D_a+D_u)]$	-

From these cases the excess probability is:

$$\Delta p^\varepsilon(B,D) = \rho(1-\rho) \cdot [(D_a-D_u)/(D_a+D_u)] \cdot [(B_a-B_u)/(B_a+B_u)] \quad (SC3)$$

Note that this last term will have to be counted twice, to account for the case in which a deleterious allele is chosen first and a beneficial second.

Now one can combine all the above terms to compute the overall excess probability of negative ε , as follows:

$$\begin{aligned}
\Delta p^\varepsilon &= \Delta p^\varepsilon(B,B) + \Delta p^\varepsilon(D,D) + 2\Delta p^\varepsilon(B,D) & (SC4) \\
&= [\rho (B_a-B_u)/(B_a+B_u)]^2 + [(1-\rho) \cdot (D_a-D_u)/(D_a+D_u)]^2 + \\
&\quad + \rho(1-\rho) \cdot [(D_a-D_u)/(D_a+D_u)] \cdot [(D_a-D_u)/(B_a+B_u)] \\
&= [\rho \cdot (B_a-B_u)/(B_a+B_u) + (1-\rho) \cdot (D_a-D_u)/(D_a+D_u)]^2
\end{aligned}$$

One first main result of this calculation is that Δp^ε can never be negative, i.e. it is always $p(\varepsilon < 0) \geq p(\varepsilon > 0)$, supporting the overall tendency towards a negative bias in the distribution of epistasis.

The above expression (equation (SC4)) describes the bias in epistasis as a function of the areas of the four regions defined by the isochange and neutrality lines in figure 2a. The ultimate goal of this derivation, however, is to know how the bias of epistasis depends on the fundamental parameters of the benefit-cost model. Hence, the next step in this derivation will consist of expressing the D_a , D_u , B_a and B_u parameters as functions of the fundamental parameters of the benefit-cost model, i.e. c_0 and b_0 . This calculation will involve simply computing the different areas based on the graph of figure 2a. In order to simplify the calculation, we will assume that both θ and λ cannot exceed a pre-defined value W . We will see later, by comparison with computer simulations, that the generality of the main conclusions is not affected by this assumption. First, from figure 2a, we notice that the following relationship holds:

$$B_a + D_a = D_u + B_u = W^2/2 \quad (SC5)$$

This implies the two following relationships:

$$B_a - B_u = D_a - D_u \quad (SC6)$$

and

$$B_a + B_u = W^2 - (D_a + D_u) \quad (SC7)$$

Therefore we can express Δp^ε as a function of ρ , $D_a - D_u$, $D_a + D_u$ and W :

$$\Delta p^\varepsilon = [\rho \cdot (D_a - D_u)[W^2 - (D_a + D_u)] + (1-\rho) \cdot (D_a - D_u)/(D_a + D_u)]^2 \quad (SC8)$$

It can be further seen from simple geometrical calculation of the trapezoidal area that:

$$D_a + D_u = [(W-1) \cdot c_0/b_0 + 1 + 1/b_0] \cdot W/2 \quad (SC9)$$

and

$$D_u - D_a = (1/b_0) - [(W-1) \cdot c_0/b_0 + 1 + 1/b_0] \cdot W/2 \quad (\text{SC10})$$

Insertion of these expressions into the above equation for Δp^ε , and with the further substitution of $b_0 = 1 + c_0$ yields, after some rearrangements:

$$\Delta p^\varepsilon = (2\rho - 1)^2 \cdot (c_0 W^2 + 2W - 2)^2 / [W^2 (Wc_0 + 2)^2] \quad (\text{SC11})$$

As one can appreciate from figure 2a that the choice $W=2$ would entail an overall balance of mutational effects around $\theta=\lambda=1$. Under this assumption, the above expression becomes:

$$\Delta p^\varepsilon = (2\rho - 1)^2 \cdot (2c_0 + 1)^2 / (2c_0 + 2)^2 \quad (\text{SC12})$$

Note the symmetry of the results for biases towards beneficial or deleterious mutations. Moreover it is $\Delta p^\varepsilon=0$ only if $\rho=1/2$. equation (SC12) is the one used to plot the $p(\varepsilon<0)$ surface of figure S2b. Note also that while this analytical derivation is based on uniform distributions in the (λ, θ) plane, our numerical calculations (See section A above) uses Gaussian sampling around (1,1). The similarity of the surfaces in figure S2a and b indicates that these differences are not substantial. This can be confirmed by solving the analytical model for alternative way of sampling points around the wild type (data not shown).

D. Analytical approximation of epistasis for a general function F

Derivation of the general epistasis expression

The general expression for epistasis shown in equation (6) in the main text can be derived using a Taylor Series approximation of the fitness f as a function (F) of two traits X and Y . Fitness for the unperturbed (wild type) system can be written as $f_0 = F(x_0, y_0)$, where x_0 and y_0 are the values of phenotypes X and Y without perturbations. Mutations i and j affect individual traits (e.g. $x_i = x_0 + \Delta x_i$, $y_i = y_0 + \Delta y_i$), as well as fitness. Fitness of the single mutants f_i and f_j can be derived from the second order Taylor Series expansion as:

$$f_i = F(x_0 + \Delta x_i, y_0 + \Delta y_i) = f_0 + \dot{F}_X \Delta x_i + \dot{F}_Y \Delta y_i + \frac{\ddot{F}_{XX} \Delta x_i^2 + \ddot{F}_{XY} \Delta x_i \Delta y_i + \ddot{F}_{YY} \Delta y_i^2}{2!} \quad (\text{SD1})$$

$$f_j = F(x_0 + \Delta x_j, y_0 + \Delta y_j) = f_0 + \dot{F}_X \Delta x_j + \dot{F}_Y \Delta y_j + \frac{\ddot{F}_{XX} \Delta x_j^2 + \ddot{F}_{XY} \Delta x_j \Delta y_j + \ddot{F}_{YY} \Delta y_j^2}{2!} \quad (\text{SD2})$$

Where $\dot{F}_X = \partial F / \partial X$, $\dot{F}_Y = \partial F / \partial Y$, $\ddot{F}_{XY} = \partial^2 F / \partial X \partial Y$ (all evaluated at (x_0, y_0)). Similarly, for the double mutant we can write:

$$f_{ij} = F(x_0 + \Delta x_{ij}, y_0 + \Delta y_{ij}) = f_0 + \dot{F}_X \Delta x_{ij} + \dot{F}_Y \Delta y_{ij} + \frac{\ddot{F}_{XX} \Delta x_{ij}^2 + \ddot{F}_{XY} \Delta x_{ij} \Delta y_{ij} + \ddot{F}_{YY} \Delta y_{ij}^2}{2!} \quad (\text{SD3})$$

We further impose zero epistasis on individual phenotype X and on phenotype Y , i.e.:

$$\begin{aligned} \frac{x_{ij}}{x_0} &= \frac{x_i}{x_0} \frac{x_j}{x_0} \rightarrow \frac{x_0 + \Delta x_{ij}}{x_0} = \frac{x_0 + \Delta x_i}{x_0} \cdot \frac{x_0 + \Delta x_j}{x_0} \\ \frac{y_{ij}}{y_0} &= \frac{y_i}{y_0} \frac{y_j}{y_0} \rightarrow \frac{y_0 + \Delta y_{ij}}{y_0} = \frac{y_0 + \Delta y_i}{y_0} \cdot \frac{y_0 + \Delta y_j}{y_0} \end{aligned} \quad (\text{SD4})$$

Using equation (SD4), Δx_{ij} and Δy_{ij} can be calculated as a function of Δx_i , Δx_j , and Δy_i , Δy_j respectively and further inserted in equation (SD3). With further substitution of $f_0=1$, the

multiplicative epistasis score ε (defined as: $\varepsilon = \frac{f_{ij}}{f_0} - \frac{f_i \cdot f_j}{f_0^2}$) becomes:

$$\begin{aligned} \varepsilon &= \Delta x_i \Delta x_j (\dot{F}_X / x_0 + \ddot{F}_{XX} - \dot{F}_X^2) \\ &+ \Delta y_i \Delta y_j (\dot{F}_Y / y_0 + \ddot{F}_{YY} - \dot{F}_Y^2) \\ &+ (\Delta x_i \Delta y_j + \Delta x_j \Delta y_i) \cdot (\ddot{F}_{XY} - \dot{F}_X \dot{F}_Y) \end{aligned} \quad (\text{SD5})$$

Equation (SD5) constitutes an approximated analytical expression for the degree of epistasis relative to a fitness $f=F(X, Y)$ between two mutations affecting multiplicatively traits X and Y .

Equation (SD5) can also be rewritten in terms of normalized perturbations, i.e. by defining $\sigma_i^X = \Delta x_i / x_0$, $\sigma_j^X = \Delta x_j / x_0$, $\sigma_i^Y = \Delta y_i / y_0$, $\sigma_j^Y = \Delta y_j / y_0$ giving rise to:

$$\begin{aligned}
\varepsilon &= \sigma_i^X \sigma_j^X (\dot{F}_X x_0 + \ddot{F}_{XX} x_0^2 - \dot{F}_X^2 x_0^2) \\
&+ \sigma_i^Y \sigma_j^Y (\dot{F}_Y y_0 + \ddot{F}_{YY} y_0^2 - \dot{F}_Y^2 y_0^2) \\
&+ (\sigma_i^X \sigma_j^Y + \sigma_j^X \sigma_i^Y) \cdot (\ddot{F}_{XY} x_0 y_0 - \dot{F}_X \dot{F}_Y x_0 y_0)
\end{aligned} \tag{SD6}$$

We use this expression to illustrate a few examples of predicted epistasis. If we apply the benefit-cost function to equation (SD6), such that $x_0=b_0, y_0=c_0, f_0 = b_0-c_0 = 1$, we can write $\sigma_i^X=\lambda_i-1, \sigma_j^X=\lambda_j-1, \sigma_i^Y=\theta_i-1, \sigma_j^Y=\theta_j-1, \dot{F}_X=1, \dot{F}_Y=-1, \ddot{F}_{XY}=\ddot{F}_{XX}=\ddot{F}_{YY}=0$ and further verify that equation (SD6) leads to equation (5) presented in the main text.

Epistasis between mutations that independently affect two traits

Non-pleiotropic mutations can be defined as mutations that affect a single trait (or module). In our model, this is represented by the conditions $\sigma_j^X=0$ and $\sigma_i^Y=0$. Upon imposing this condition to equation (SD6), its first two lines disappear and equation (SD6) becomes $\varepsilon = x_0 y_0 \sigma_i^X \sigma_j^Y (\ddot{F}_{XY} - \dot{F}_X \dot{F}_Y)$. Since x_0, y_0, σ_i^X and σ_j^Y are nonzero, the condition of zero-epistasis for mutations affecting independent modules would be $\ddot{F}_{XY} - \dot{F}_X \dot{F}_Y = 0$, which is equivalent to $\partial^2 \log F / (\partial x \partial y) = 0$, at (x_0, y_0) :

$$\begin{aligned}
\left. \frac{\partial^2 \log F}{\partial X \partial Y} \right|_{x_0, y_0} &= \left(\frac{\partial}{\partial Y} \left(\frac{1}{F} \frac{\partial F}{\partial X} \right) \right) \bigg|_{x_0, y_0} = \left(-\frac{1}{F^2} \frac{\partial F}{\partial Y} \frac{\partial F}{\partial X} + \frac{1}{F} \frac{\partial^2 F}{\partial X \partial Y} \right) \bigg|_{x_0, y_0} \\
&= \left(\frac{\partial^2 F}{\partial X \partial Y} - \frac{\partial^2 F}{\partial X \partial Y} \right) \bigg|_{x_0, y_0} \quad (\because F(x_0, y_0) = 1) \\
&= \ddot{F}_{XY} - \dot{F}_X \dot{F}_Y \text{ evaluated at } (x_0, y_0)
\end{aligned}$$

Examples of specific functional dependencies

Here we provide several examples of epistasis for different biologically relevant F functions. In the following expressions, for simplicity, we use x and y in place of x_0 and y_0 . The calculations are based on equation SD6, and utilize the condition that $F(x_0, y_0)=1$.

$$(1) F(X, Y) = X^n Y^m$$

(For the special case $n=m=1$, this would estimate, for example, epistasis between the enzyme level and k_{cat} in the Michaelis-Menten equation)

$$\begin{aligned}
\varepsilon &= \sigma_i^X \sigma_j^X (-n^2 x^n y^m (x^n y^m - 1)) + \sigma_i^Y \sigma_j^Y (m^2 (1 - x^{2n} y^{2m})) + (\sigma_i^X \sigma_j^Y + \sigma_j^X \sigma_i^Y) * 0 \\
&= 0 + 0 + 0 = 0
\end{aligned}$$

$$(2) F(X, Y) = G(X) * H(Y) = \exp(X) * \exp(Y)$$

(This dependence exemplifies a nontrivial case of factorizable F function, i.e. a case in which $G(X)$ and $H(Y)$ are not simple power functions)

$$\varepsilon = \sigma_i^X \sigma_j^X(x) + \sigma_i^Y \sigma_j^Y(y) + (\sigma_i^X \sigma_j^Y + \sigma_j^X \sigma_i^Y) * 0 = \sigma_i^X \sigma_j^X(x) + \sigma_i^Y \sigma_j^Y(y)$$

$\varepsilon = 0$ if allele i and j act on distinct traits with no pleiotropic effect ($\sigma_j^X=0$ and $\sigma_i^Y=0$).

$$(3) F = \frac{a + bX}{c + dY}$$

(This equation could be applied to estimating epistasis between V_{\max} and K_m in the Michaelis-Menten equation)

$$\begin{aligned} \varepsilon &= \sigma_i^X \sigma_j^X \left(\frac{abx}{(a + bx)^2} \right) + \sigma_i^Y \sigma_j^Y \left(-\frac{cdy}{(c + dy)^2} \right) + (\sigma_i^X \sigma_j^Y + \sigma_j^X \sigma_i^Y) * 0 \\ &= \sigma_i^X \sigma_j^X \left(\frac{abx}{(a + bx)^2} \right) + \sigma_i^Y \sigma_j^Y \left(-\frac{cdy}{(c + dy)^2} \right) \end{aligned}$$

For Michaelis-Menten parameters: $X=V_{\max}$, $Y=K_m$, ($a = 0$, $b = c = S$ is constant, $d=1$)

$$\varepsilon = \sigma_i^X \sigma_j^X(0) + \sigma_i^Y \sigma_j^Y \left(-\frac{SK_m}{(S + K_m)^2} \right) + (\sigma_i^X \sigma_j^Y + \sigma_j^X \sigma_i^Y) * 0 = -\sigma_i^Y \sigma_j^Y \left(\frac{SK_m}{(S + K_m)^2} \right)$$

$\varepsilon = 0$ if allele i and j act on distinct traits with no pleiotropic effect ($\sigma_i^Y=0$).

$$(4) F(X,Y) = aX^n + bY^m$$

(This example could be seen as a generalized version of the benefit-cost model)

$$\varepsilon = \sigma_i^X \sigma_j^X (-an^2 x^n (ax^n - 1)) + \sigma_i^Y \sigma_j^Y (-bm^2 y^m (by^m - 1)) + \sigma_i^X \sigma_j^Y + \sigma_j^X \sigma_i^Y (-abnm x^n y^m) * 0$$

In general $\varepsilon \neq 0$ in such case.