FISHER'S GEOMETRIC MODEL OF ADAPTATION MEETS THE FUNCTIONAL SYNTHESIS: DATA ON PAIRWISE EPISTASIS FOR FITNESS YIELDS INSIGHTS INTO THE SHAPE AND SIZE OF PHENOTYPE SPACE

Daniel M. Weinreich[1,2] and Jennifer L. Knies[1,3]

[1]Department of Ecology and Evolutionary Biology, and Center for Computational Molecular Biology, Brown University, Providence, Rhode Island 02912

[2]E-mail: Daniel_Weinreich@Brown.edu

[3]Current address: Department of Molecular Biology and Chemistry, Christopher Newport University, Newport News, Virginia 23606

The functional synthesis uses experimental methods from molecular biology, biochemistry and structural biology to decompose evolutionarily important mutations into their more proximal mechanistic determinants. However these methods are technically challenging and expensive. Noting strong formal parallels between R.A. Fisher's geometric model of adaptation and a recent model for the phenotypic basis of protein evolution, we sought to use the former to make inferences into the latter using data on pairwise fitness epistasis between mutations. We present an analytic framework for classifying pairs of mutations with respect to similarity of underlying mechanism on this basis, and also show that these data can yield an estimate of the number of mutationally labile phenotypes underlying fitness effects. We use computer simulations to explore the robustness of our approach to violations of analytic assumptions and analyze several recently published datasets. This work provides a theoretical complement to the functional synthesis as well as a novel test of Fisher's geometric model.

KEY WORDS: Adaptation, epistasis, fitness, models/simulations, mutations, pleiotropy.

A mutation's effect on phenotype commonly depends on the presence or absence of other mutations in the genome, a phenomenon called epistasis. This word was coined 100 years ago by William Bateson (1909) to describe departures from the Mendelian 9:3:3:1 ratios expected in the F_2 generation of a dihybrid cross involving unlinked loci (Phillips 1998, 2008). The principled analyses of such departures can yield insights into the mechanisms that determine phenotype (Avery and Wasserman 1992; Griffiths et al. 2002; Roth et al. 2009). For example in the blue-eyed Mary (*Collinsia parviflora*), crossing true-breeding plants with white and magenta flowers yields 100% blue flowers in the F_1, and blue, magenta, and white flowers at ratios 9:3:4 in the F_2. Thus, a mutation at one locus is masked in organisms homozygous mutant at the other, implying that the gene product of the masked locus acts downstream of the gene product of the masking locus.

Mendelian phenotypes are discrete, but epistasis is also widespread in continuous phenotypes and has important implications for human medicine, agriculture, and evolutionary biology (Mackay 2001). Population genetics traditionally define epistasis for fitness between two mutations i and j as

$$\epsilon_{ij} = W_0 W_{ij} - W_i W_j, \qquad (1)$$

where W_x is the absolute fitness of an organism carrying mutation(s) x and W_0 is the fitness of the ancestor. We further partition

epistasis ($\varepsilon \neq 0$) into positive and negative epistasis (meaning that the fitness of the double-mutant is higher and lower, respectively, than expected).

The principled analyses of the sign of fitness epistasis between deleterious mutations in pairs of genes can also yield insights into the topology of an organism's metabolic network. Building on the metabolic control analysis of Kascer and Burns (1973), Szathmáry (1993) showed that mutations in pairs of genes whose products act in the same linear pathway should exhibit positive epistasis, whereas mutations in pairs of genes whose products act in parallel, redundant pathways should exhibit negative epistasis (Segrè et al. 2005; D. M. Weinreich, unpubl. ms.). Recently, this reasoning has been used with remarkable success to correctly reconstruct metabolic networks from microbial growth rate data for pairwise gene deletions estimated computationally (Segrè et al. 2005) or measured experimentally (Costanzo et al. 2010; He et al. 2010).

Although metabolic control analysis suggests that deleterious mutations in the same gene should only exhibit negative fitness epistasis (Szathmáry 1993), experimental work illustrates that the situation is more complicated (e.g., Lunzer et al. 2005a; Weinreich et al. 2006; Bridgham et al. 2007; Lozevsky et al. 2009). Indeed, these results have stimulated experimental interest in developing a quantitative understanding of how mechanistically more proximal phenotypes drive protein evolution (the functional synthesis of Dean and Thornton 2007). We therefore sought a theoretical perspective complementary to this recent empirical work. We were specifically motivated by the fact that in many systems high-throughput reverse genetics and fitness assays are now practical, whereas the dissection of the underlying molecular biology of fitness effects remains more challenging. Our approach yields insight into more proximal determinants of fitness using only data on pairwise epistasis between deleterious mutations.

We begin here from the observation (Weinreich 2010) that a principled model of protein evolution first proposed by De-Pristo et al. (2005; see also Camps et al. 2007; Gu 2007; Tokuriki et al. 2008) bears strong formal similarities to Fisher's geometric model (Fisher 1930; hereafter the FGM). The FGM maps from an abstract representation of the space of all conceivable organismal phenotypes to fitness, and recently Martin et al. (Martin and Lenormand 2006; Martin et al. 2007; Chevin et al. 2010) have developed quantitative predictions for patterns of pairwise fitness epistasis under the premise of this abstract phenotypic space. We inverted this logic by asking whether data on fitness epistasis might allow us to quantitatively characterize the actual phenotypic space of the FGM in specific experimental systems. Note too that although motivated by a model of protein evolution, our approach can be applied at any level of organismal organization.

We first present a novel analytic framework that uses epistatic interactions to characterize pairs of deleterious mutations with re-spect to similarity of phenotypic effect in the FGM. In addition, given data on pairwise epistatic interactions among some set of deleterious mutations, our framework yields an estimate of the total number of mutationally labile dimensions in the phenotypic space of the corresponding FGM. We then use computer simulations to explore the robustness of results to violations of analytic assumptions. Finally, we examine several published datasets. Taken together this study offers a novel theoretical complement to ongoing efforts at dissecting the phenotypic determinants of fitness as well as a quantitatively rigorous examination of Fisher's geometric model.

PROTEIN EVOLUTION AND FISHER'S GEOMETRIC MODEL

DePristo et al. (2005) assert three facts about protein biology. First, that in addition to affecting functional properties (e.g., catalytic activity of an enzyme or binding specificity of a transcription factor), mutations in protein coding genes influence organismal fitness via several other molecular phenotypes. These may include gene transcription and message translation rates, the protein's native-form folding kinetics and thermodynamics, and its misfolding, aggregation, and degradation rates. Second, many of these phenotypes are commonly under stabilizing selection, that is intermediate values are optimal. For example, although too little folding stability might render most copies of an enzyme nonfunctional, too much stability will also hinder function (Malcolm et al. 1990) as some flexibility, termed "breathing," is essential. (Zhou et al. 1998; Wang et al. 2002; Tomatis et al. 2005) (But see Bloom et al. 2005; Zeldovich et al. 2007 for an alternative perspective.) Finally, many mutations act pleiotropically within a protein (e.g., Raquet et al. 1995; Sideraki et al. 2001; Wang et al. 2002; Tomatis et al. 2005). Thus, DePristo et al. (2005) argue that protein evolution is influenced by stabilizing selection acting on multiple phenotypes responding pleiotropically to mutation. Those authors note that fitness epistasis emerges as a consequence of the nonlinear mapping from phenotype to fitness inherent in stabilizing selection (see also Brodie 2000; Martin et al. 2007).

These facts are illustrated in Figure 1 with data for two missense mutations in the TEM-1 β-lactamase allele, responsible for bacterial resistance against β-lactam drugs such as penicillin and cephalosporins. Introducing the G238S mutation (representing serine for glycine at residue 238; amino acid numbering as in Ambler et al. 1991) into the TEM-1 allele increases resistance against the drug cefotaxime (Wang et al. 2002; Weinreich et al. 2006). *In vitro* this mutation increases catalytic activity against this drug (Raquet et al. 1994; Sideraki et al. 2001; Wang et al. 2002), but it reduces thermodynamic stability (Raquet et al. 1995; Wang et al. 2002). In contrast, the M182T mutation reduces cefotaxime resistance when introduced into TEM-1 but increases resistance in the presence of G238S (Wang et al. 2002;
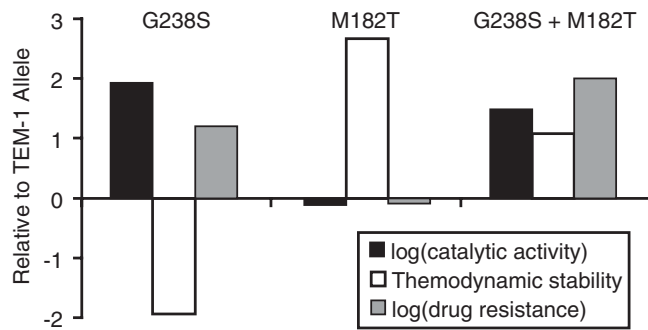
Figure 1. Effect of two β-lactamase mutations on cefotaxime drug resistance (minimum inhibitory concentration, in μg/mL), catalytic activity (k_{cat}/K_M in $M^{-1} \cdot s^{-1}$), and thermodynamic folding stability (ΔG in kcal/mol). Values relative to the TEM-1 wild-type allele. M182T only conditionally improves cefotaxime resistance: it reduces drug resistance (gray bars) on the TEM-1 background but increases it sharply in the presence of G238S. (Data from Wang et al. 2002; Weinreich et al. 2006; see Weinreich et al. 2005 for the evolutionary implications of such epistasis.) In contrast, note that for both mutations the effect on catalytic activity and thermodynamic folding stability is roughly additive. Figure adapted from Weinreich (2010).

Weinreich et al. 2006). M182T reduces *in vitro* activity while increasing stability (Huang and Palzkill 1997). Thus, interplay between the pleiotropy of these mutations is thought to render M182T conditionally beneficial because the double mutant enjoys increased hydrolysis (Wang et al. 2002) without loss of stability (Huang and Palzkill 1997). (See Weinreich et al. 2005, 2006 for the evolutionary significance of conditionally beneficial or "sign epistatic" mutations.)

Fisher's geometric model (Fig. 2A) begins from the premise that the fundamental attribute of adaptation are the many simultaneous "features of conformity" (Fisher 1930, p. 38), both within an organism and between the organism and its environment. Fisher first imagined a continuous, multidimensional phenotype space; any organism can be represented by some point in this space. The biologically optimal combination of phenotypes is at the origin, and fitness elsewhere is a declining function of the distance to the origin. Thus, stabilizing selection acts simultaneously on multiple phenotypes. Second, Fisher assumed that mutations displace an organism in an arbitrary direction in phenotype space; i.e., mutations usually act pleiotropically. Thus, the FGM captures the essential elements of protein evolution outlined above (Weinreich 2010), and this point is illustrated qualitatively in Figure 2B for the G238S and M182T mutations of the TEM-1 allele of β-lactamase. The FGM is one of the few phenotypic models of evolution (reviewed in Orr 2005a) and has received extensive theoretical (Kimura 1983; Hartl and Taubes 1996; Hartl and Taubes 1998; Orr 1998, 1999, 2005b; 2006; Poon and Otto 2000; Welch and Waxman 2003; Waxman and Welch 2005; Martin and Lenor-

mand 2006; Waxman 2006; Sella 2009; Chevin et al. 2010; Le Nagard et al. 2011; Sellis et al. 2011) and experimental (Burch and Chao 1999; Martin et al. 2007; Tenaillon et al. 2007; MacLean et al. 2010; Rokyta et al. 2011) attention.

Together these facts motivate this study, which explores pairwise epistasis for fitness between deleterious mutations under the FGM. Specifically any two mutational vectors in phenotype space define an angle θ between them, and θ influences their fitness epistasis as illustrated in Figure 3A and B. We now invert this logic to make inferences about the geometry of the phenotypic space defined by the FGM from data on epistasis for fitness between deleterious mutations.

Methods
### COMPUTING PHENOTYPIC ANGLE Θ BETWEEN MUTATIONS UNDER FISHER'S GEOMETRIC MODEL

We begin by making two assumptions: that fitness is a Gaussian function of phenotypic distance from the optimum, and that phenotypic space is a vector space, that is that mutations are additive in phenotype space (e.g., Lande 1980; Martin and Lenormand 2006; Martin et al. 2007; but see computer simulations later). The first assumption approximates many functions close to the optimum (Lande 1980; Waxman and Welch 2005; Martin and Lenormand 2006). The second assumption is equivalent to supposing that epistasis acts only on fitness and not on underlying phenotypes. There is good empirical support for this assumption among missense mutations in proteins (Lunzer et al. 2005a; Zeldovich et al. 2007; although see Yokoyama et al. 2008 for an important counterexample) and moreover recent theoretical results were found to be robust to its violation (Martin et al. 2007).

Following Martin et al. (2007; see also Fig. 2 here), given n phenotypes under selection, we represent each organism by a column vector z whose ith component ($1 \leq i \leq n$) gives the organism's deviation at the ith phenotype from the optimal value. (Thus, the optimal phenotype z_{opt} is the column vector of all zeros.) Similarly, we represent mutations by column vector dz, whose ith component gives the mutation's perturbation of the ith phenotype. Writing the ancestral organism's phenotype z_0, the absolute fitness of an offspring carrying mutation dz is then

$$W(z_0+dz) = \text{Exp}\left[-\frac{1}{2}(z_0+dz)^T S(z_0+dz)\right], \quad (2a)$$

where superscript T represents transposition and S is the positive semidefinite $n \times n$ selective covariance matrix. For notational convenience, we shall also sometimes employ log-transformed absolute fitness

$$w(z_0+dz) = \ln[W(z_0+dz)] = -\frac{1}{2}(z_0+dz)^T S(z_0+dz). \quad (2b)$$
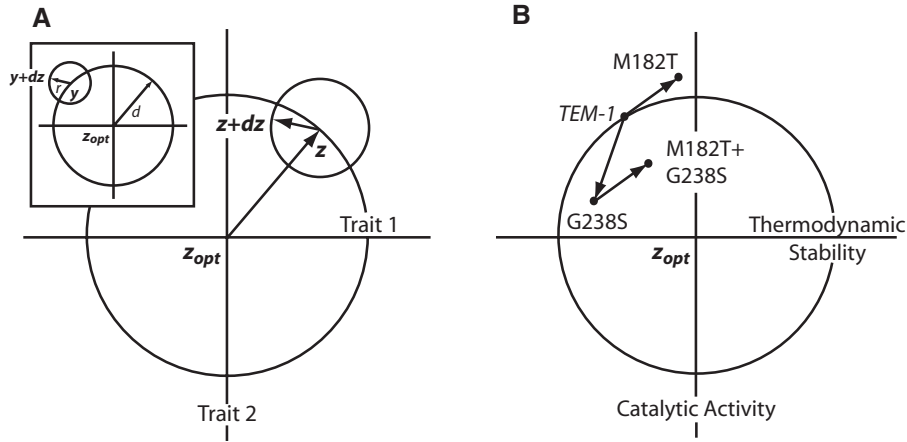
Figure 2. Fisher's geometric model of adaptation for $n = 2$ phenotypes. **(A)** All combinations of phenotype values are represented in n-dimensional space (here, the plane); the optimal value is represented by the point labeled z_{opt} and an individual-labeled z is displaced from this optimum (as perhaps following an environmental perturbation). To be beneficial, a mutation dz on wild-type z must yield a phenotype $z + dz$ lying within the circle passing through z and centered at z_{opt}. Inset: y represents another organism with fitness equal to that of z. Note, however, that mutation dz is seen to exhibit epistasis: it is beneficial on z but deleterious on y. **(B)** Mutational interactions in TEM-1 β-lactamase between catalytic activity and thermodynamic stability in determining cefotaxime resistance (see Fig. 1), represented qualitatively in Fisher's geometric model. Figure adapted from Weinreich (2010).

Now writing the weighted inner product $\langle dz_i, dz_j \rangle = dz_i^T S dz_j$ and corresponding vector norm $\|dz_i\|^2 = \langle dz_i, dz_i \rangle = dz_i^T S dz_i$, the phenotypic angle between two mutational vectors dz_i and dz_j is defined by

$$\cos(\theta) = \frac{\langle dz_i, dz_j \rangle}{\sqrt{\|dz_i\|} \cdot \sqrt{\|dz_j\|}} = \frac{dz_i^T S dz_j}{\sqrt{dz_i^T S dz_i} \cdot \sqrt{dz_j^T S dz_j}}, \quad (3)$$

the right-hand side of which can in some cases be written in terms of empirical fitness data. To see this note first that under equation (2b), $w(z_0 + dz_i + dz_j) + w(z_0) - w(z_0 + dz_i) - w(z_0 + dz_j) = -dz_i^T S dz_j$. This expression captures the epistasis between mutations dz_i and dz_j, and while slightly different than the classical definition (our equation (1), it is identical to e_{ij} of Martin et al. (2007; see their eq. 2, although that derivation is given in terms of relative rather than absolute fitnesses; see also Bonhoeffer et al. 2004). Importantly, assuming a Gaussian fitness function this empirical expression for $dz_i^T S dz_j$ is correct for all ancestral phenotypes z_0 because it simply reflects the (invariant) curvature of the function, as already noted (Martin et al. 2007). In contrast the denominator of equation (3) can be expressed in terms of empirical fitness data only by assuming $z_0 = z_{opt}$. In this case, we invert our equation (2b) to find $2w(dz_i) = -dz_i^T S dz_i$. Substituting these results into equation (3) yields

$$\theta_{FGM} =$$
$$\arccos\left\{-\frac{w(z_0 + dz_i + dz_j) + w(z_0) - w(z_0 + dz_i) - w(z_0 + dz_j)}{\sqrt{-2w(dz_i)}\sqrt{-2w(dz_j)}}\right\} \quad (4)$$

as the phenotypic angle between mutations dz_i and dz_j.

Equation (4) thus uses fitness data to give insight into the phenotypic space defined by the FGM; we write θ_{FGM} to highlight this connection. Specifically, the cosine of θ_{FGM} is Martin et al. (2007)'s epistasis e_{ij} between mutations i and j, normalized by the magnitude of their individual effects. Consequently pairs of strongly epistatic mutations are parallel or antiparallel in phenotype space (i.e., have correlated phenotypic effects, although the converse is not true in general; see Results).

Importantly, our norm in phenotypic space must obey the triangle and reverse triangle inequalities:

$$\text{Max}(\|dz_i\|, \|dz_j\|) - \text{Min}(\|dz_i\|, \|dz_j\|) \leq \|dz_i + dz_j\|$$
$$\leq \|dz_i\| + \|dz_j\|. \quad (5)$$

These bounds correspond respectively to $\theta = 180°$ and $\theta = 0°$ and are shown as heavy lines in Figure 3C. Violating this assumption (here as a consequence of experimental error in fitness assays or violations of assumptions leading to eq. 4) causes the expression inside the curly brackets of equation (4) to fall outside the interval $[-1, 1]$ and will yield imaginary values of θ_{FGM}.

However as illustrated by Figure 3D, all real and imaginary values of θ_{FGM} lie on a continuum: values of θ_{FGM} of the form $0° + ci$ are simply more extreme than $\theta = 0°$ and values of θ_{FGM} of the form $180° + ci$ are simply more extreme than $\theta = 180°$. (Here i is the imaginary unit defined by $(-1)^{1/2}$ and c is some real constant.) We therefore treat such observations as categorical data lying on either side of the continuous interval $[0°, 180°]$.
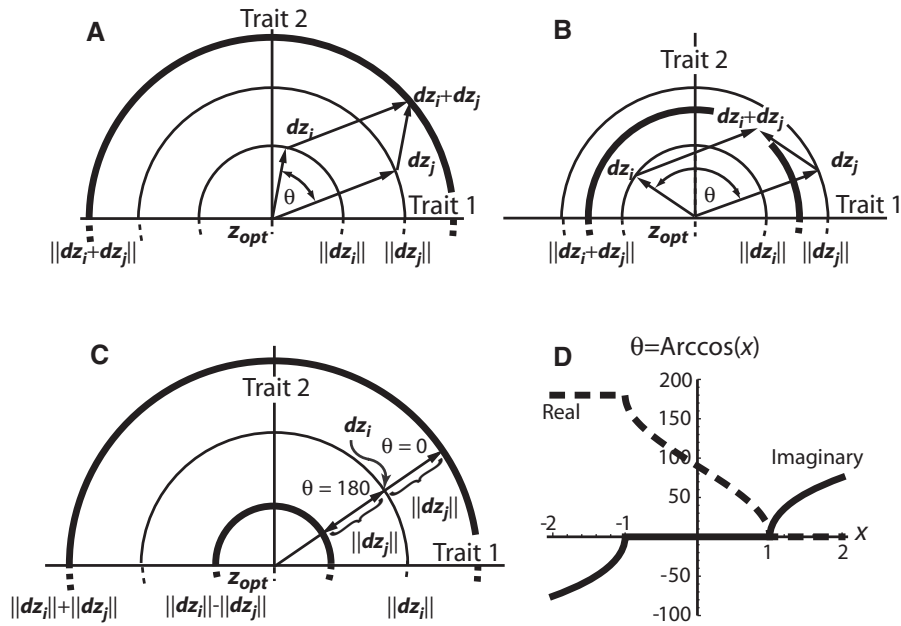
Figure 3. Phenotypic angle θ between mutations dz_i and dz_j under Fisher's geometric model. Placing any two mutational vectors dz_i and dz_j on a common ancestral phenotype defines a plane and an angle θ between vectors. (A, B) $\|dz_i\|$ and $\|dz_j\|$ are both constant, where $\|x\|$ is the norm of vector x defined as $x^T Sx$. Thin circles have radii $\|dz_i\|$ and $\|dz_j\|$; heavy circles have radii $\|dz_i+dz_j\|$. Note dependence of $\|dz_i+dz_j\|$ (and hence, of epistasis) on θ. (C) Maximum and minimum values of $\|dz_i+dz_j\|$ allowed under the model (eq. 5) correspond respectively to $\theta = 0°$ and $180°$. Experimental error in fitness assays and violations of model assumptions can yield imaginary estimates of θ. Specifically when $Max(\|dz_i\|, \|dz_j\|) - Min(\|dz_i\|, \|dz_j\|) > \|dz_i+dz_j\|$, the phenotypic vector representing dz_i+dz_j is inferred to lie inside the inner heavy circle and the arccos($\cdot$) in equation (4) will yield θ_{FGM} values of the form $180° + ci$. Conversely, when $\|dz_i+dz_j\| > \|dz_i\| + \|dz_j\|$, this vector is inferred to lie outside the outer heavy circle and θ_{FGM} will be of the form $0° + ci$. (D) Real (dashed) and imaginary (solid) parts of $\theta = arccos(x) \cdot 180°/\pi$. Note that in any case we can view results as lying on a single continuum: as already suggested in C: values of θ_{FGM} of the form $0° + ci$ lie beyond $\theta = 0°$ and values of θ_{FGM} of the form $180° + ci$ lie beyond $\theta = 180°$. See Results and Discussion for further details.

COMPUTING THE NUMBER OF PHENOTYPIC DIMENSIONS FROM FITNESS DATA

The distribution of values of θ_{FGM} observed also contains information about n, the number of mutationally labile dimensions in the underlying phenotype space. It has previously been shown that that the probability density function for angles θ between some reference direction and a set of vectors of random orientation in an n-dimensional space is

$$f(\theta) = Z \sin(\theta)^{n-2} \qquad (6)$$

(Hartl and Taubes 1998; Poon and Otto 2000, where Z is the scaling constant required to get a proper probability density function over $[0°, 180°]$). This is also the probability density function for θ between vectors sampled from a multivariate Gaussian distribution with mean 0 and covariance matrix $M = c \cdot I$, where c is any positive scalar and I is the identity matrix (not shown). Thus, assuming this sampling scheme, equation (6) gives the expected distribution of θ_{FGM} between some focal mutation dz_i and all others in the dataset. But because this distribution is independent of the choice of focal mutation, it applies equally to the unconditioned distribution of θ_{FGM} across an entire dataset. After discarding imaginary values of θ_{FGM} we thus employed this density function in a likelihood framework to estimate $\hat{n}$ together with 95% confidence intervals (CIs).

COMPUTER SIMULATION

Computer simulations were run using MATLAB R2011a (Mathworks, Natick, MA) to confirm analytical results and to explore the effect of experimental noise in fitness assays and relaxations of model assumptions. All code has been archived at https://github.com/weinreich/FGM-FunctionalSynthesis. Following Martin et al. (2006, 2007), we assume that the mutational probability density function is Gaussian with covariance matrix M. For the purpose of exploring the robustness of equation (4) we further assume a rotationally symmetric mutational covariance matrix $M = m \times I$ (because its derivation is independent of M) with standard deviation $m = 0.1$ phenotypic units. We relaxed this assumption to assess our ability to estimate phenotypic dimensionality. Because equation (4) is also independent of selection matrix S, we set $S = I$ throughout.

SENSITIVITY OF ESTIMATES OF PAIRWISE PHENOTYPIC ANGLE θ_{FGM}

In each run, we sampled 200 random mutational vectors dz_i and dz_j and computed θ_{true} between each of the resulting $\binom{200}{2} = 19,900$. pairs of mutations using equation (3) and $S = I$. We then tabulated the mean and standard deviation for corresponding values of θ_{FGM} (as described next) within a sliding window of θ_{true} values $10°$ wide. We also tabulated the frequency of each kind of imaginary value of θ_{FGM} observed in each sliding window.

To explore the sensitivity of equation (4) to experimental noise in fitness assays, we first added a normal deviate with mean zero and standard deviation σ_{noise} to fitness equation (2a). For each pair of mutations, we then computed θ_{FGM} with equation (4) using the average of R replicate noisy fitness measures for each mutational vector. We explored values of experimental noise $\sigma_{noise} = 0.01$ and 0.02, experimental replication $R = 10$ and 40, and phenotypic dimensions, $n = 2, 5$, and 20. Because fitness values in our simulations are in the range of 0.8–1.0 (not shown), values of σ_{noise} here are roughly comparable to the CV values shown in Table 1 (experimental standard error normalize by experimental mean values, see Analysis of Empirical Data). Simulated levels of replication are consistent with current experimental protocols (e.g., Sanjuán et al. 2004), and simulated phenotypic dimensionality values are consistent with estimates here (Fig. 6) and elsewhere (Tenaillon et al. 2007).

To test consequences of our assumption that the ancestral phenotype is at the fitness optimum, we also ran simulations in which $W(z_0) = 0.99$ and 0.90 for phenotypic dimension $n = 2$, 5, and 20. We accomplished this by setting each component of z_0 to $(2\ln[W(z_0)]/n)^{1/2}$. To relax the assumption that mutations are additive in phenotype space, we added a phenotypic epistasis vector E to each double mutant before applying equation (2a). Each component of E was drawn from a normal distribution with mean 0 and standard deviation $\sigma_{epistasis}$, parameterized as a proportion of mean mutational vector length m. Values of $\sigma_{epistasis} = 0.1m$ and $0.5m$ were examined.

SENSITIVITY OF ESTIMATES OF PHENOTYPIC DIMENSIONALITY n

To explore the sensitivity of our approach to estimating phenotypic dimensionality n in the face of experimental noise and violations of model assumptions, we simulated samples of 10 or 25 mutations each, setting $n = 2, 5$, or 20, and allowing experimental noise in fitness assays, nonoptimal ancestral phenotypes, or phenotypic epistasis as described earlier. We also examined violations of the assumption of rotational symmetry in mutational matrix M underlying equation (6). Here again, samples of 10 or 25 mutations were generated in $n = 2, 5$, or 20 dimensions with mutational matrix $M = m \times I$ as above, except that we set matrix entries $m_{1,1} = m/2$, or $m_{1,2} = m/2$, or both.

For each sample, we discarded imaginary values of θ_{FGM}, calculated the likelihood surface for a range of values of n between 2 and 40 in steps of 0.05, and recorded the maximum likelihood estimate together with the 95% CI bounds. These simulated datasets are comparable in size (10 and 25 mutations yield at most 45 and 300 comparisons, respectively) and phenotypic dimensionality to those in Table 3.

DATA ANALYSIS

Theoretical results developed here were applied to the seven published datasets shown in Table 1. Only deleterious mutations described in Elena and Lenski (1997a,b) and Sanjuán (2004) were considered, and in all datasets lethal mutations were removed (as required by eq. 4; for this reason we analyzed MIC values presented in Lozevsky et al. 2009 rather than IC_{50} values). We then used equation (4) to compute θ_{FGM} between all pairs of remaining mutations in each dataset, and the number of dimensions of the underlying phenotypic space was estimated as described. Finally, we used the Kolmogorov–Smirnov goodness-of-fit test (Sokal and Rohlf 1995) as implemented in Matlab to compare the observed distributions of θ_{FGM} for each dataset with model expectations (eq. 6) for given $\hat{n}$.

Results

Under analytic assumptions, Fisher's geometric model allows us to use epistasis for fitness between pairs of deleterious mutations to characterize those pairs with respect to similarity of underlying mechanism. Specifically, pairs of mutations with very similar mechanisms will exhibit strong epistasis and values of θ_{FGM} near $0°$ or $180°$, although the converse need not be true. (As the numerator of eq. 3 illustrates, epistasis will only be absent if the two mutational vectors are orthogonal in the space defined by matrix S, but such mutations will only exhibit uncorrelated phenotypes if all off-diagonal elements of S are 0.) Moreover, the distribution of epistasis across a panel of deleterious mutations can yield an estimate of the number of mutationally labile phenotypes contributing to fitness.

ESTIMATING θ_{FGM} WHEN ANALYTIC ASSUMPTIONS ARE RELAXED

Under analytic assumptions (i.e., setting $\sigma_{noise} = \sigma_{epistasis} = 0$, $z_0 = z_{opt}$ and random positive semidefinite S) simulation results were precisely concordant with equation (4) for all parameter values examined (not shown). Figures 4 and 5 present simulation results that explore sensitivity of equation (4) to experimental noise and to relaxing model assumptions, respectively. Each panel displays the mean real θ_{FGM} values ($\pm$ 1 SD) across a sliding window of

Table 1. Published biological datasets analyzed.

Organism, gene(s)	Number of variable sites or loci analyzed	Source of mutations	Fitness assay	Fitness assay CV	Citation
Escherichia coli, isopropylmalate dehydrogenase	6	Site-directed	Chemostat competition assay against the wild-type	0.0037[1]	Lunzer et al. (2005a, b)
Escherichia coli, β-lactamase	5	Site-directed	Minimum inhibitory concentration against cefotaxime	0.087	Weinreich et al. (2006)
Plasmodium falciparum dihydrofolate reductase	4	Site-directed	Minimum inhibitory concentration against pyrimethamine	No data reported	Lozovsky et al. (2009)
Escherichia coli, gyrA, rpoB, and rpsL	3	Random	Paired growth assays against the wild-type in LB	0.68	Trindade et al. (2009a, b)
Eschericia coli, Mini-Tn10 mutants	8	Random	Paired growth assay against ancestor in DM25	0.018	Elena and Lenski (1997a, b)
Vesicular stomatitis virus N, P, M, G, and L	28	Random	Intrinsic growth rate in mammalian cell culture	0.056	Sanjuan et al. (2004)
Aspergillus niger, arg, pyr, leu, phe, oli, and crn	7	Random	Radial growth rate of colony	0.030	de Visser et al. (2009)

[1]Reported in Lunzer et al. (2002).

θ_{true} values 10° wide. In addition, we show the frequency of each of the two classes of imaginary values of θ_{FGM} observed each in each sliding window. Finally, in each panel, we report the r^2 value for a linear regression of real estimates of θ forced through the origin (see Discussion).

We note first that in all cases, imaginary values of θ_{FGM} of the form $0° + ci$ are primarily observed when θ_{true} is small, whereas imaginary values of the form $180° + ci$ are associated with large values of θ_{true}. This is to be expected given our proposed interpretation of such observations (Fig. 3D). Moreover, conditioned on θ_{FGM} being real, θ_{FGM} almost always overestimates θ_{true} when θ_{true} is small and underestimates it when large. This is a simple consequence of our conditioning: the "missing," more extreme estimates of θ_{FGM} are to be found in the imaginary realizations. Finally, in almost all cases summing across all values of θ_{true}, the frequency of imaginary values of θ_{FGM} of the form $0° + ci$ exceeds the frequency of values of the form $180° + ci$. This can be understood by examining Figure 3C, which illustrates that for all parameter values the area outside the larger heavy circle [representing $\|dz_i\| + \|dz_j\|$] exceeds that inside the smaller heavy circle [representing $\text{Max}(\|dz_i\|, \|dz_j\|) - \text{Min}(\|dz_i\|, \|dz_j\|)$]. Put another way, experimental noise and violations of model assumptions perturb our estimate of where the vector $dz_i + dz_j$ lies. But when θ_{true} is near 180° such perturbation is less

likely to result in an imaginary estimate of θ_{FGM} than when it is near 0°.

Figure 4 illustrates that conditioning on the observation of real values of θ_{FGM}, equation (4) can estimate θ_{true} quite well from modestly noisy fitness data. Strong, nearly unbiased predictive power is observed in the first panel and standard deviation declines as phenotypic dimension increases (the trend continues for $n = 20$; not shown). This can be understood as a consequence of two facts. First, as dimension increases the probability distribution of realized θ_{FGM} becomes increasingly concentrated around 90° (Hartl and Taubes 1996; Poon and Otto 2000; see also Figs. 5, 6 here). This alone accounts for the drop in frequency of imaginary values of θ_{FGM} as dimension rises. Second, because $S = I$ in simulations, as θ_{FGM} approaches 90° $W(dz_i + dz_j)$ approaches $W(dz_i)W(dz_i)$. Thus, even after allowing noise in fitness assays, the numerator of the expression inside curly braces in equation (4) will tend to be small, minimizing the standard deviation in θ_{FGM}. Increasing experimental error (σ_{noise}) increases standard deviation of θ_{FGM} for all values of θ, although as expected this effect can be offset by additional fitness assay replicates (R).

Sensitivity of equation (4) to violations of two assumptions is explored in Figure 5. Figure 5A illustrates that when the ancestor is not at the phenotypic optimum (i.e., $z_{0-} \neq z_{opt}$), values of θ_{FGM} are skewed toward intermediate values. As noted earlier, this effect

Table 2. Maximum likelihood estimate and confidence intervals on phenotypic dimensionality $\hat{n}$ in computer simulations.

True phenotypic dimensionality, n	2	5	20	2	5	20
Number of mutations sampled	10	10	10	25	25	25
Using default values[1]	1.95	4.10	17.65	2.00	5.05	18.70
	(1.70 – 2.35)	(3.45 – 6.10)	(12.20 – 24.00)	(1.90 – 2.10)	(4.50 – 5.65)	(16.15 – 21.70)
$\sigma_{noise} = 0.01, R = 10$	2.55	4.75	25.05	2.70	5.10	20.80
	(2.05 – 3.20)	(3.60 – 6.45)	(16.80 – 29.60)	(2.50 – 2.95)	(4.60 – 5.75)	(17.90 – 23.95)
$\sigma_{noise} = 0.02, R = 10$	2.70	5.35	20.75	2.50	4.95	19.15
	(2.20 – 3.50)	(4.05 – 7.35)	(14.00 – 24.65)	(2.30 – 2.75)	(4.45 – 5.55)	(16.55 – 22.20)
$\sigma_{noise} = 0.02, R = 40$	2.35	5.45	15.90	2.30	4.45	20.15
	(1.95 – 2.95)	(4.10 – 7.50)	(11.05 – 23.20)	(2.15 -2.50)	(4.00 – 4.95)	(17.40 – 23.35)
$W(z_0) = 0.99$	2.85	6.15	21.80	3.65	5.10	20.60
	(2.25 – 3.70)	(4.55 – 8.50)	(14.95 – 29.05)	(3.30 – 4.10)	(4.55 – 5.70)	(17.75 – 23.90)
$W(z_0) = 0.90$	7.00	12.50	25.20	5.60	7.75	21.90
	(5.00 – 10.15)	(8.80 – 18.20)	(17.20 – 33.85)	(4.90 – 6.45)	(6.80 – 8.90)	(18.90 – 24.45)
$\sigma_{epistasis} = 0.1m$	2.15	4.45	15.85	2.20	5.00	19.30
	(1.85 – 2.60)	(3.40 – 6.00)	(11.05 – 23.15)	(2.05 – 2.35)	(4.45 – 5.65)	(16.70 – 22.40)
$\sigma_{epistasis} = 0.5m$	2.50	5.55	25.20	2.75	4.85	19.55
	(2.05 – 3.20)	(4.15 – 7.65)	(17.20 – 33.85)	(2.55 – 3.05)	(4.35 – 5.45)	(16.90 – 22.70)
$m_{1,1} = m/2$	1.85	5.05	16.05	1.90	4.45	18.40
	(1.60 – 2.15)	(3.85 – 6.90)	(11.15 – 23.40)	(1.80 – 2.00)	(4.00 – 5.00)	(15.90 – 21.35)
$m_{1,2} = m_{2,1} = m/2$	1.60	3.20	16.15	1.65	3.85	16.55
	(1.45 – 1.85)	(2.60 – 4.15)	(11.25 – 23.60)	(1.60 – 1.75)	(3.50 – 4.30)	(14.30 – 19.15)
$m_{1,1} = m_{1,2} = m_{2,1} = m/2$	1.55	5.20	20.35	1.70	4.20	17.45
	(1.40 – 1.75)	(3.90 – 7.10)	(14.05 – 29.90)	(1.65 – 1.80)	(3.80 – 4.70)	(15.10 – 20.25)

[1]Default values used throughout except where noted. σ_{noise} = experimental noise (default: 0.0); R = experimental replicates (default: 1), $W(z_0)$ = fitness of ancestor (default: 1.0), $\sigma_{epistasis}$ = phenotypic epistasis (default: 0.0m), $M = m \times I$ mutational covariance matrix (default: m 0.1), selection covariance matrix $S = I$ throughout.

is entirely the result of errors introduced through the denominator of equation (4) because $-dz_i^T S dz_j$ is only equal to $2\ln[W(dz_i)]$ when $z_{0.} = z_{opt}$. Consequently, the magnitude of the effect is seen to increase as the fitness of z_0 declines. However, the effect is moderated with increasing phenotypic dimension because the phenotypic perturbation introduced by any given fitness drop in z_0 is proportional to $n^{-\frac{1}{2}}$.

Figure 5B illustrates the consequences of phenotypic epistasis. Again, increased deviations from model assumptions increase the error in equation (4). We also note that in contrast to the situation in Figure 5A, increasing phenotypic dimensions render results more sensitive to violations of model assumptions. This is to be expected as we add a normally distributed random deviate to each component of the double-mutant's phenotype. Thus, the net error in $z_0 + dz_i + dz_j$ is proportional to $n^{\frac{1}{2}}$.

ESTIMATING PHENOTYPIC DIMENSIONALITY n WHEN ANALYTIC ASSUMPTIONS ARE RELAXED

Sensitivity of our estimator for phenotypic dimensionality in the face of experimental noise and violation of model assumptions is

shown in Table 2. In general, we note that the maximum likelihood estimate is a bit below the center of the 95% CI, signaling a slightly positive skew to the likelihood function. We also note a slight bias toward overestimating phenotypic dimensionality in the face of experimental error, most notably when $z_0 \neq z_{opt}$. This is to be expected: by conditioning on real values of θ_{FGM} we are discarding realizations that otherwise would have fallen in the tails of the distribution of angles (Fig. 3), making phenotypic dimensionality appear higher than it is (eq. 6). However, deviations from rotationally symmetric mutations bias dimensionality estimates downward. This is also to be expected: as mutational variance in one dimension is reduced (here, $m_{1,1}$ has been reduced from 0.1 to 0.05) mutations will become increasingly skewed in the remaining dimension(s). And similarly, as mutational covariance is increased (here $m_{1,2} = m_{2,1}$ has increased from 0.0 to 0.05), mutations will more often have a common orientation.

ANALYSIS OF EMPIRICAL DATA

We computed θ_{FGM} between all pairs of mutations in each of the seven published datasets shown in Table 1, and used those

Table 3. Analyses of published datasets.[1]

Organism, gene(s)	Number of comparisons	Number of times $\theta_{FGM} = 0 + ci$ observed	Number of times $\theta_{FGM} = 180 + ci$ observed	Kolmogorov–Smirnov test statistic and associated P-value	Mean θ_{FGM}
Escherichia coli, isopropylmalate dehydrogenase	67	43	0	0.502, 4.1×10^{-6}	58.4°
Escherichia coli, β-lactamase	10	1	0	0.27, 0.46	94.8°
Plasmodium falciparum dihydrofolate reductase	6	1	1	0.59, 0.07	133.0°
Escherichia coli, gyrA, rpoB, and rpsL	98	0	16	0.39, 8.4×10^{-12}	117.4°
Escherichia coli, Mini-Tn10 mutants	21	6	4	0.37, 0.07	104.5°
Vesicular stomatitis virus	31	1	5	0.122, 0.80	92.7°
Aspergillus niger, arg, pyr, leu, phe, oli, and crn	14	0	0	0.652, 2.4×10^{-6}	120.1°
Escherichia coli, isopropylmalate dehydrogenase, bottom-up θ_{FS}[2]	67	N/A	N/A	0.422, 4.4×10^{-11}	77.4°

[1]See Table 1 for citations.
[2]Bottom-up estimates based on angles between phenotypic vectors. See text for details.

distributions to compute the estimated phenotypic dimensionality $\hat{n}$ together with the 95% CI, as described. Observed and expected distributions, $\hat{n}$ values and confidence intervals, and P-values for goodness-of-fit to equation (6) are all shown for each dataset in Figure 6. Further details of these analyses are provided in Table 3.

Discussion

Evolution by natural selection enriches populations for those organisms with higher reproductive success (Darwin 1859), but what is the underlying biology responsible for these differences? Fitness is a complex property that emerges from many mechanistically more proximal phenotypes, and R.A. Fisher's geometric model (1930; here, the FGM) provides a popular, explicit, and abstract representations of that mapping (although see Kimura and Maruyama 1966; Kondrashov 1988; Rice 2000 for other mappings). Noting the close qualitative parallels to a recent model of protein evolution (DePristo et al. 2005; see Fig. 2 here), we thought to use the FGM to explore our ability to make inferences into this underlying phenotypic space from data on epistasis for fitness between deleterious mutations. We were motivated by the many low-throughput (e.g., Lunzer et al. 2005a; Weinreich et al. 2006; Lozevsky et al. 2009) and high-throughput (e.g., Costanzo et al. 2010; He et al. 2010) datasets now being developed. We sought two practical contributions for the experimentalist: the

ability to group mechanistically related mutations (i.e., those with highly correlated phenotypic effects) and the ability to make gross statements about the phenotypic complexity of the system (i.e., the dimensionality of phenotypic space).

INTERPRETING VALUES OF θ_{FGM}

Importantly however, our approach is complicated by the fact that equation (4) can return biologically ambiguous results in the form of imaginary values (Table 3). Indeed, all datasets considered here except one (de Visser et al. 2009) include some imaginary values of θ_{FGM}; these range in frequency from 10% (Weinreich et al. 2006) to 64% (Lunzer et al. 2005a). In simulations, we observed that imaginary values occur as a consequence of measurement error in fitness assays (Fig. 4) as well as violations of model assumptions (Fig. 5). However, no one of these factors appears to fully explain the source of these observations in the data.

Simulations demonstrate that the fraction of imaginary values should rise with experimental error in fitness assays and drop with phenotypic dimensionality (Fig. 4). We thus first tabulated experimental error (CV) for six of the seven datasets examined (Table 1; no error estimates are given in Lozevsky et al. 2009), but contrary to the suggestion from simulation results, the fraction of imaginary values of θ_{FGM} observed is weakly and negatively correlated with experimental CV ($r^2 = 7.3\%$). We next regressed the fraction of imaginary values of θ_{FGM} against our
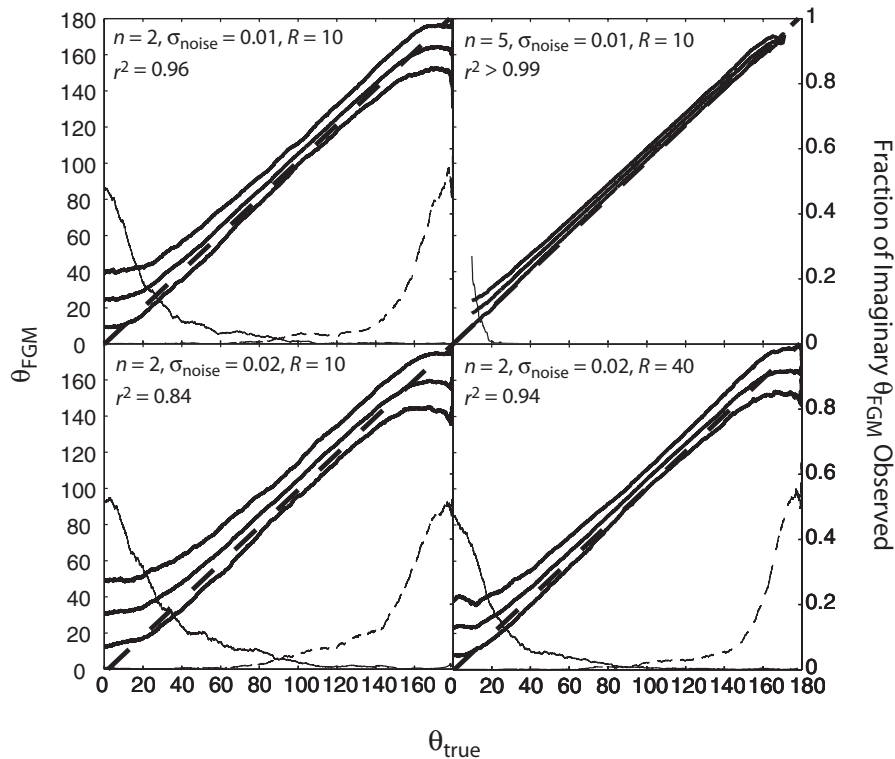
Figure 4. Estimation of θ_{FGM} in computer simulations allowing experimental error in fitness assay. Equation (4) can return real and imaginary values. Heavy lines represent mean ±1 SD of real values of θ_{FGM} (eq. 4) for a given θ_{true} (computed with eq. 3) in a sliding window 10° wide; heavy dashed line represents $\theta_{FGM} = \theta_{true}$. (Scale on left y-axis.) Light lines represent fraction of imaginary values of θ_{FGM}'s in the same sliding window (solid: frequency of imaginary θ_{FGM} with real part = 0°; dashed: frequency of imaginary θ_{FGM}'s with real part = 180°; scale on right y-axis). All parameter values as shown (phenotypic dimensionality n, experimental noise σ_{noise}, experimental replicates R). Phenotypic epistasis ($\sigma_{epistasis}$) equal to 0, $z_0 = z_{opt}$, and $S = I$.

estimated phenotypic dimensionality for each dataset. Although this correlation is also negative, consistent with expectation, it is also only weakly so ($r^2 = 17.7\%$). Thus, although some imaginary values of θ_{FGM} may be the consequence of experimental error in fitness assays, we wondered whether they may not also derive in part from violations of model assumptions: $z_0 = z_{opt}$, additivity among phenotypic vectors, and a Gaussian fitness function.

For three datasets (Lunzer et al. 2005a; de Visser et al. 2009; Trindade et al. 2009bb), we computed epistasis on a naturally occurring wild-type isolate, which we might reasonably expect to be at or near the (local) fitness optimum, and these studies describe no beneficial mutations on those backgrounds. Two datasets (Weinreich et al. 2006; Lozevsky et al. 2009) employ drug resistance as a proxy for fitness, and we computed epistasis on the highest-resistance variant described. Again, these studies describe no beneficial mutations on those backgrounds. (And in one case extensive mutagenesis has failed to reveal a higher-resistance variant; Orencia et al. 2001; Salverda et al. 2011.) Although beneficial mutations on the ancestor were described in the last two datasets (Elena and Lenski 1997b; Sanjuán et al. 2004) we believe it is

unlikely that $z_0 \neq z_{opt}$ is responsible for many of the imaginary values of θ_{FGM} in Table 3. Indeed, in the case where the starting genotype is demonstrably far the fitness optimum (Sanjuán et al. 2004), the frequency of imaginary values of θ_{FGM} is comparatively modest.

To our knowledge, data on phenotypic additivity are scarce. One case however comes from Lunzer et al. (2005a,b). In addition to measuring organismal fitness, those authors performed *in vitro* assays for four functionally proximal biochemical phenotypes likely responsible for fitness effects of mutations in their system. They find that mutational effects on log-transformed values of these phenotypes are very nearly additive ($r^2 \geq 0.92$). Thus at present, we find little evidence to suggest that phenotypes are not additive.

Finally, we acknowledge that non-Gaussian fitness functions can also give rise to imaginary values of θ_{FGM}. Although sensitivity of our approach to this effect has not been explored, principled models of the phenotype-to-fitness function for both IMDH (eq. 3 in Lunzer et al. 2005a) and β-lactamase (eq. 4 in Brown et al. 2009) are distinctly non-Gaussian. Following others (e.g., Perelson and Oster 1979; Lapedes and Farber 2001) we are presently
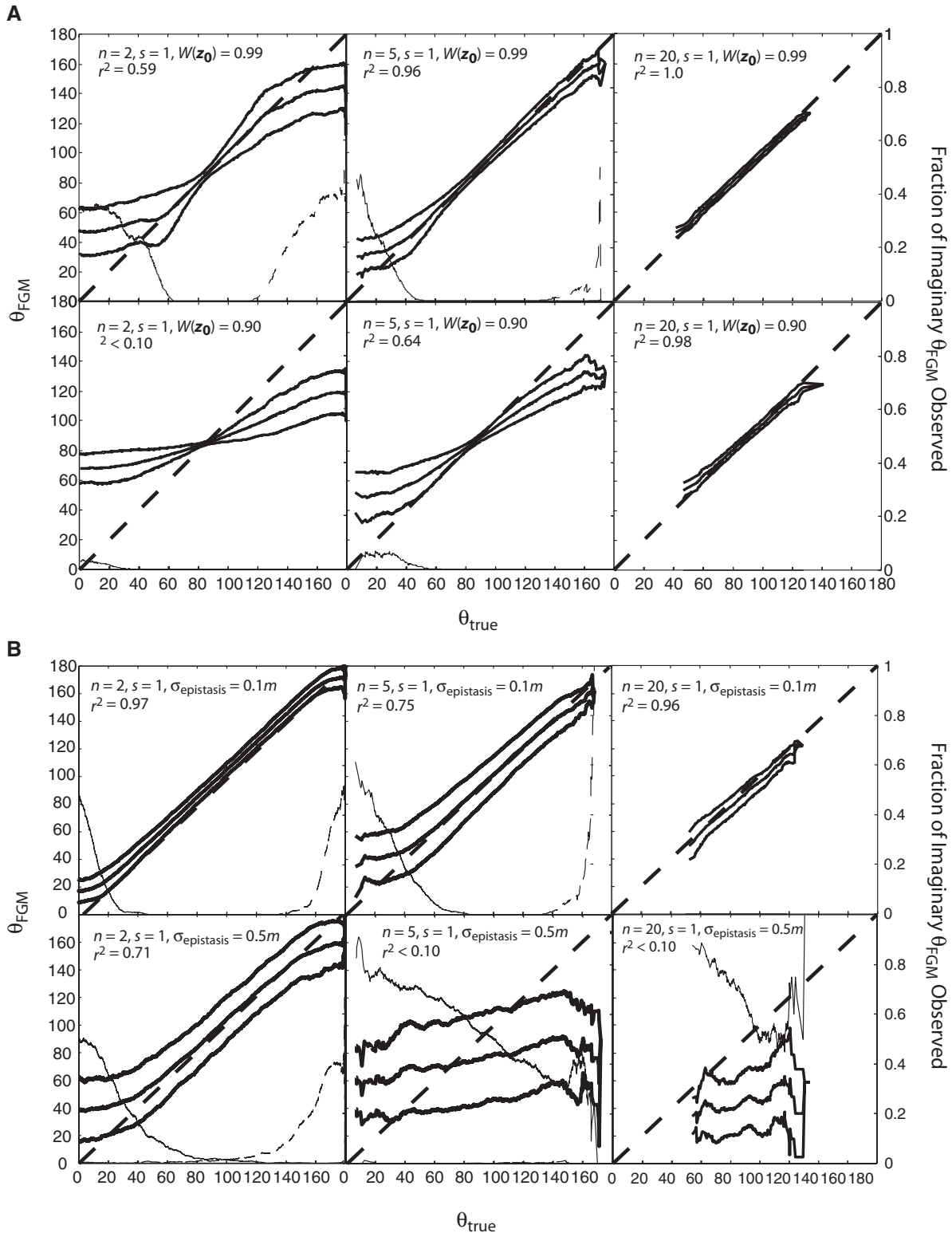
Figure 5. Estimation of θ_{FGM} in computer simulations when model assumptions are relaxed. (A) Relaxing $z_0 = z_{opt}$. (B) Relaxing phenotypic additivity. Note that as n increases, values of θ_{true} become increasingly concentrated around 90° (Hartl and Taubes 1998; Poon and Otto 2000). See Figure 4 legend for additional details.
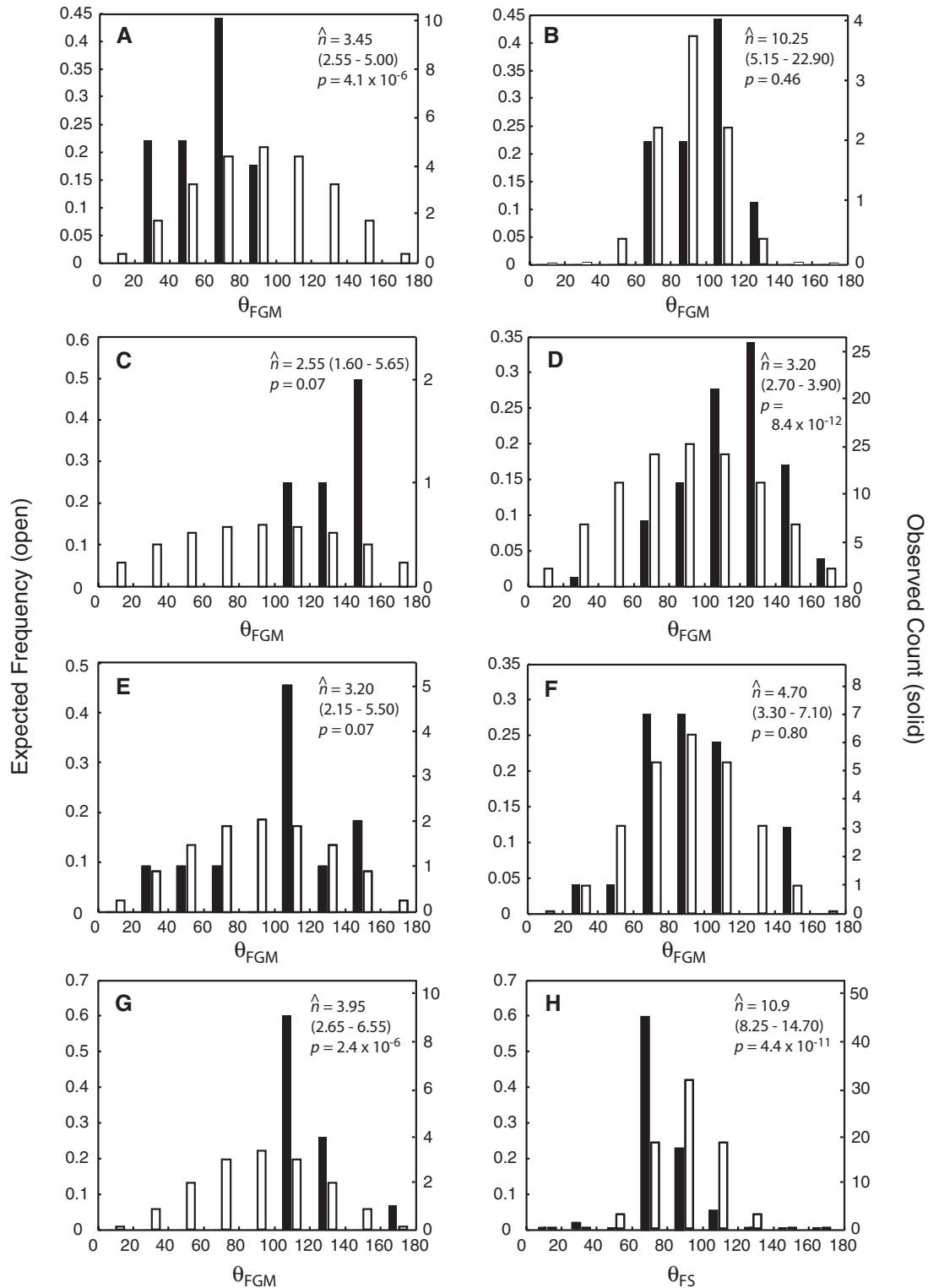
Figure 6. Frequency of real pairwise angles between pairs of deleterious mutations in published datasets shown in Table 1. Filled: θ_{FGM} computed with equation (4). Open: expectation as described, corresponding to phenotypic dimensionality $\hat{n}$ estimated for each dataset. (A) Lunzer et al. (2005a, b), (B) Weinreich et al. (2006), (C) Lozovsky et al. (2009), (D) Trindade et al. (2009a, b), (E) Elena and Lenski(1997a, b), (F) Sanjuan et al (2004), (G) de Visser et al. (2009), (H) Bottom-up estimates of θ_{FS} derived from Lunzer et al. (2005b). (See text for details.)
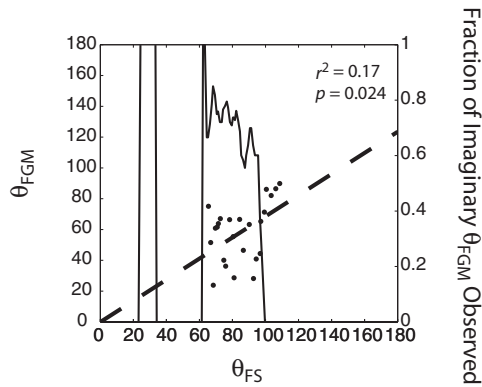
Figure 7. Bottom-up analysis of phenotypic angles among mutations in isopropylmalate dehydrogenase examined in Lunzer et al. (2005a, b). For each pair of mutations, x–axis represents pairwise phenotypic angle θ_{FS} computed with equation (3) using phenotypic perturbations in log-transformed cofactor performance and binding affinity estimated by Lunzer et al. (2005a, b). Left y-axis represents real θ_{FGM} values computed from fitness data using equation (4); dashed line shows linear regression through the origin. Right y-axis represents the fraction of imaginary values of θ_{FGM} in a sliding window 10° wide. Only imaginary θ_{FGM} values of the form $0 + ci$ were observed in this dataset. Cf. Figures 4 and 5.

pursuing an alternative approach that relieves us of the burden of this assumption.

TOP-DOWN VERSUS BOTTOM-UP ESTIMATES OF θ

Here we have employed what may be regarded as a top-down approach to quantitatively characterize the phenotypic space underlying the FGM. The functional synthesis (Dean and Thornton 2007) represents a complementary bottom-up approach: characterization of the mechanistically proximal phenotypic traits responsible for fitness effects of particular mutations. As already noted, in addition to fitness Lunzer et al. (2005a, b) characterized mutational perturbation in four candidate proximal phenotypes related to enzyme–substrate affinity and enzyme kinetics. These data allow further examination of equation (4), because we can explicitly represent mutations as phenotypic vectors ($\Delta \ln[K_m^{NAD}]$, $\Delta \ln[K_m^{NADP}]$, $\Delta \ln[V_{max}^{NAD}/K_m^{NAD}]$, $\Delta \ln[V_{max}^{NADP}/K_m^{NADP}]$), where Δ is relative to the allele of highest fitness. Application of equation (3) to these vectors thus yields independent measures of pairwise angles θ between mutations in phenotypic space. We designate these values θ_{FS} to reflect their connection to the functional synthesis.

To determine how well θ_{FS} predicts θ_{FGM}, we regressed the latter on the former (Fig. 7). Consistent with expectations we observe a significantly positive slope between real values of θ_{FGM} and θ_{FS} (Spearman rank correlation coefficient $\rho = 0.408$, $P = 0.024$ by a permutation test). The slope of the regression through

the origin (0.69) is different from 1.0, although this should perhaps not surprise us because the numeric angle θ between two vectors is sensitive to scaling of phenotypic space. For example, Lunzer et al. (2005a) report values of V_{max}/K_M in units of $M^{-1} \cdot s^{-1}$; rescaling V_{max}/K_M in units of $mM^{-1} \cdot s^{-1}$ would change all values of θ_{FS} without changing any of the biology. Indeed, we are substantially heartened by this result for two reasons. First, rescaling dimensions of phenotypic space can even perturb rank orders of angles when $n > 2$. Second, although θ_{FS} offers direct access to mutational vectors in phenotype space, θ_{FGM} values are influenced by the (unknown) asymmetries of the true selection covariance matrix S. Figure 7 also shows that the frequency of imaginary values of θ_{FGM} of the form $0° + ci$ declines as θ_{FS} rises, consistent with model expectations. (No values of θ_{FGM} of the form $180° + c$ were observed in this dataset; Table 3.)

INTERPRETING VALUES OF *n*

Our second objective was to provide some access to the phenotypic complexity underlying fitness. Interestingly, values reported here (Fig. 6) fall between those previously developed under the FGM by others (0.2–2.5 in Martin and Lenormand 2006; and 10–45 in Tenaillon et al. 2007; see also Lourenço et al. 2011 for an attempt to bridge those divergent estimates). Although equation (4) examines curvature of the fitness function at the fitness optimum, Martin and Lenormand (2006) employ data from mutation accumulation experiments and their results thus characterize the fitness function over a wider mutational scope. Finally, estimates in Tenaillon et al. (2007) are based on differences in fitness at mutation/selection/drift equilibrium as a function of population size, and may therefore represent the analysis covering the largest genetic scale.

As above, phenotypic data in Lunzer et al. (2005b) also allow us to perform a bottom-up complement to our top-down estimation of $\hat{n}$. The distribution of θ_{FS} (Fig. 6H) implies that these vectors lie in a higher dimensionality space than do results based on θ_{FGM} (Fig. 6A; $\hat{n} = 10.90$ vs. 3.45), although the two distributions are statistically indistinguishable ($G = 2.45$, $P > 0.98$, df = 9).

A bias in $\hat{n}$ derives from our assumption that mutational vectors are uniformly distributed in phenotypic space. To the extent that they are not, ours are underestimates (Table 2). We thus regard our approach as estimating an "effective" number of dimensions in phenotype space (i.e., the number of dimensions of equal and uncorrelated mutational lability that would give rise to the given distribution of θ_{FGM}). Although an analytic analog to equation (6) for arbitrary mutational covariance matrix M appears possible, it would not be helpful here because we would then be faced with estimating both $\hat{n}$ and $\hat{M}$ from single distributions of θ_{FGM}.

In addition, because equation (6) is undefined when θ is imaginary, it was necessary to omit such values of θ_{FGM} when

we estimated phenotypic dimensionality. This protocol biases $\hat{n}$ upward. Recognizing the diversity of datasets in our study, we however observe no signal of this bias: the correlation between $\hat{n}$ and the frequency of imaginary values of θ_{FGM} is in fact weakly negative.

Finally, the distribution of θ_{FGM} deviates significantly from the expectation given by equation (6) in three cases (Fig. 6; Table 3). In one (Lunzer et al. 2005a, b), the mean (58.4°) is far less than expectation (as is the mean θ_{FS} from the same system: 77.4°), implying that these mutations are underdispersed in phenotype space. This effect is already visible in Figure 3 of Lunzer et al. (2005a), which shows that all phenotypic mutations lie in a narrow sector emanating from the wild type. Means are larger than expectation in the other two cases (de Visser et al. 2009; Trindade et al. 2009b; 120.2° and 117.4°, respectively), implying mutational overdispersion, and we imagine that in these systems ridges of high mutational density radiate from the wild type, separated by valleys of low mutational density approximately 120° wide. Both of these situations imply correlations between mutations (positive in the first case and negative in the latter two; note that these correlations are distinct from correlations within mutations represented by the matrix ***M***), although the biological and evolutionary implications are as yet unclear to us. No particular association between these deviations and the sources of mutations (Table 1) was observed.

CONCLUSIONS

Using Fisher's geometric model we have developed a novel approach to characterizing aspects of the phenotypic determinants of deleterious mutations using data on pairwise epistasis. Despite our original motivation for adopting the FGM (Fig. 2), we are generally heartened by the apparent utility of our approach when applied to the heterogeneous datasets (Table 1: four from bacteria, two from eukaryotes and one from a virus). Moreover, although growth rate is the fitness assay in five cases, drug resistance is employed in two. And finally, three of the datasets consider mutations in a single gene while the other four use mutations more broadly distributed across the genome.

We conclude by noting that much of evolutionary biology rightly focuses on beneficial mutations, those which fuel adaptation. In contrast, we have only considered epistasis between deleterious mutations, both because they are far more plentiful and because our algebraic framework cannot accommodate beneficial mutations. Moreover statistical patterns of fitness epistasis among deleterious mutations do not always mirror those among beneficial mutations (Sanjuán et al. 2004; Rokyta et al. 2011). However, our interest in epistasis here is limited to making inferences into the mechanistically proximal phenotypes underlying mutational effects on fitness. And as it seems reasonable to assume that deleterious and beneficial mutations operate in the same phenotypic space, we suggest that application of the framework developed here will yield dividends equally applicable to the molecular basis of Darwinian evolution.

LITERATURE CITED

Ambler, R. P., A. F. W. Coulson, J.-M. Frère, J.-M. Ghuysen, B. Joris, M. Forsman, R. C. Levesque, G. Tiraby, and S. G. Waley. 1991. A standard numbering scheme for the Class A β-lactamases. Biochem. J. 276:269–272.

Avery, L., and S. Wasserman. 1992. Ordering gene function: the interpretation of epistasis in regulatory hierarchies. Trends Genet. 8:312–316.

Bateson, W. 1909. Mendel's principles of heredity. Cambridge Univ. Press, Cambridge, U.K.

Bloom, J. D., J. J. Silberg, C. O. Wilke, D. A. Drummond, C. Adami, and F. H. Arnold. 2005. Thermodynamic prediction of protein neutrality. Proc. Natl. Acad. Sci. USA 102:606–611.

Bonhoeffer, S., C. Chappey, N. T. Parkin, J. M. Whitcomb, and C. J. Petropolous. 2004. Evidence for positive epistasis in HIV-1. Science 306:1547–1550.

Bridgham, J. T., S. M. Carroll, and J. W. Thornton. 2007. Evolution of hormone-receptor complexity by molecular exploitation. Science 312:97–100.

Brodie, E. D., III. 2000. Why evolutionary genetics does not always add up. Pp. 3–20 in J. B. Wolf, E. D. Brodie, III, and M. J. Wade, eds. Epistasis and the evolutionary process. Oxford Univ. Press, New York.

Brown, K. M., M. A. DePristo, D. M. Weinreich, and D. L. Hartl. 2009. Temporal constraints on the incorporation of regulatory mutants in evolutionary pathways. Mol. Biol. Evol. 26:2455–2462.

Burch, C. L., and L. Chao. 1999. Evolution by small steps and rugged landscapes in the RNA virus ø6. Genetics 151:921–927.

Camps, M., A. Herman, E. Loh, and L. A. Loeb. 2007. Genetic constraints on protein evolution. Crit. Rev. Biochem. Mol. Biol. 42:313–326.

Chevin, L.-M., G. Martin, and T. Lenormand. 2010. Fisher's model and the genomics of adaptation: restricted pleiotropy, heterogeneous mutation, and parallel evolution. Evolution 64:3213–3231.

Costanzo, M., A. Baryshnikova, J. Bellay, Y. Kim, E. D. Spear, C. S. Sevier, H. Ding, J. L. Y. Koh, K. Toufighi, S. Mostafavi, et al. 2010. The genetic landscape of a cell. Science 327:425–431.

Darwin, C. 1859. On the origin of species by means of natural selection. John Murray, Lond.

de Visser, J. A. G. M., S.-C. Park, and J. Krug. 2009. Exploring the effect of sex on empirical fitness landscapes. Am. Nat. 174:S15–S30.

Dean, A. M., and J. W. Thornton. 2007. Mechanistic approaches to the study of evolution: the functional synthesis. Nat. Rev. Genet. 8:675–688.

DePristo, M. A., D. M. Weinreich, and D. L. Hartl. 2005. Missense meandering through sequence space: a biophysical perspective on protein evolution. Nat. Rev. Genet. 6:678–687.

Elena, S. F., and R. E. Lenski. 1997a. Data from: test of synergistic interactions among deleterious mutations in bacteria. doi:10.5061/dryad.rg8mb.

———. 1997b. Test of synergistic interactions among deleterious mutations in bacteria. Nature 390:395–397.

Fisher, R. A. 1930. The genetical theory of natural selection. Clarendon Press, Oxford, U.K.

Griffiths, A. J. F., W. M. Gelbart, R. C. Lewontin, and J. Miller. 2002. Modern genetic analysis: integrating genes and chromosomes. W.H. Freeman, New York.

Gu, X. 2007. Evolutionary framework for protein sequence evolution and gene pleiotropy. Genetics 175:1813–1822.

Hartl, D. L., and C. H. Taubes. 1996. Compensatory nearly neutral mutations: selection without adaptation. J. Theor. Biol. 182:303–309.

———. 1998. Toward a theory of evolutionary adaptation. Genetica 102/103:525–533.

He, X., W. Qian, Z. Wang, Y. Li, and J. Zhang. 2010. Prevalent positive epistasis in *Escherichia coli* and *Saccharomyces cerevisiae* metabolic networks. Nature Genet. 42:272–276.

Huang, W., and T. Palzkill. 1997. A natural polymorphism in β-lactamase is a global suppressor. Proc. Natl. Acad. Sci. USA 94:8801–8806.

Kacser, H., and J. Burns. 1973. The control of flux. Symp. Soc. Exp. Biol. 27:65–104.

Kimura, M. 1983. The neutral theory of molecular evolution. Cambridge Univ. Press, Cambridge, U.K.

Kimura, M., and T. Maruyama. 1966. The mutational load with epistatic gene interactions in fitness. Genetics 54:1337–1351.

Kondrashov, A. S. 1988. Deleterious mutations and the evolution of sex. Nature 336:435–440.

Lande, R. 1980. The genetic covariance between characters maintained by pleiotropic mutations. Genetics 94:203–215.

Lapedes, A., and R. Farber. 2001. The geometric of shape space: application to influenza. J. Theor. Biol. 212:57–69.

Le Nagard, H., L. Chao, and O. Tenaillon. 2011. The emergence of complexity and restricted pleiotropy in adapting networks. BMC Evol. Biol. 11:326.

Lourenço, J., N. Galtier, and S. Glémin. 2011. Complexity, pleiotropy, and the fitness effect of mutations. Evolution 65:1559–1571.

Lozevsky, E. R., T. Chookajorn, K. M. Brown, M. Imwong, P. J. Shaw, S. Kamchonwongpaisan, D. E. Neafsey, D. M. Weinreich, and D. L. Hartl. 2009. Stepwise acquisition of pyrimethamine resistance in the malaria parasite. Proc. Natl. Acad. Sci. 106:12025–12030.

Lunzer, M., A. Natarajan, D. E. Dykuisen, and A. M. Dean. 2002. Enzyme kinetics, substitutable resources and competition: from biochemistry to frequency-dependent selection in *lac*. Genetics 162:485–499.

Lunzer, M., S. P. Miller, R. Felsheim, and A. M. Dean. 2005a. The biochemical architecture of an ancient adaptive landscape. Science 310:499–501.

———. 2005b. Data from: the biochemical architecture of an ancient adaptive landscape. doi:10.5061/dryad.7nd70.

Mackay, T. F. C. 2001. Quantitative trait loci in *Drosophila*. Nat. Rev. Genet. 2:11–20.

MacLean, R. C., G. G. Perron, and A. Gardner. 2010. Diminishing returns from beneficial mutations and pervasive epistasis shape the fitness landscape for Rifampicin resistance in *Pseudomonas aeruginosa*. Genetics 186:1345–1354.

Malcolm, B. A., K. P. Wilson, B. W. Matthews, J. F. Kirsch, and A. C. Wilson. 1990. Ancestral lysozymes reconstructed, neutrality tested, and thermostability linked to hydrocarbon packing. Nature 345:86–89.

Martin, G., and T. Lenormand. 2006. A general multivariate extension of Fisher's geometric model and the distribution of fitness effects across species. Evolution 60:893–907.

Martin, G., S. F. Elena, and T. Lenormand. 2007. Distributions of epistasis in microbes fit predictions from a fitness landscape model. Nat. Genet. 39:555–560.

Orencia, M. C., J. S. Yoon, J. E. Ness, W. P. C. Stemmer, and R. D. Stevens. 2001. Predicting the emergence of antibiotic resistance by directed evolution and structural analysis. Nat. Struct. Biol. 8:238–242.

Orr, H. A. 1998. The population genetics of adaptation: the distribution of factors fixed during adaptive evolution. Evolution 52:935–949.

———. 1999. The evolutionary genetics of adaptation: a simulation study. Genet Res 74:207–214.

———. 2005a. The genetic theory of adaptation: a brief history. Nature Rev. Genet. 6:119–127.

———. 2005b. Theories of adaptation: what they do and don't say. Genetica 123:3–13.

———. 2006. The distribution of fitness effects among beneficial mutations in Fisher's geometric model of adaptation. J Theor Biol 238:279–285.

Perelson, A. S., and G. F. Oster. 1979. Theoretical studies of clonal selection: minimal antibody repertoire size and reliability of self-non-self discrimination. J. Theor. Biol. 81:645–670.

Phillips, P. C. 1998. The language of gene interaction. Genetics 149:1167–1171.

———. 2008. Epistasis—the essential role of gene interactions in the structure and evolution of genetic systems. Nat. Rev. Genet. 9:855–867.

Poon, A., and S. P. Otto. 2000. Compensating for our load of mutations: freezing the meltdown of small populations. Evolution 54:1467–1479.

Raquet, X., J. Lamotte-Brasseur, E. Fonzé, S. Goussard, P. Courvalin, and J.-M. Frère. 1994. TEM β-lactamase mutants hydrolysing third-generation cephalosporins. J. Mol. Biol. 2444:625–639.

Raquet, X., M. Vanhove, J. Lamotte-Brasseur, S. Goussard, P. Courvalin, and J.-M. Frère. 1995. Stability of TEM β-lactamase mutants hydrolyzing third generation cephalosporins. Proteins 23:63–72.

Rice, S. H. 2000. The evolution of developmental interactions. Pp. 82–98 *in* J. B. Wolf, E. D. I. Brodie and M. J. Wade, eds. Epistasis and the evolutionary process. Oxford Univ. Press, Oxford, U.K.

Rokyta, D. R., P. Joyce, S. B. Caudle, C. Miller, C. J. Beisel, and H. A. Wichman. 2011. Epistasis between beneficial mutations and the phenotype-to-fitness map for a ssdna virus. PLoS Genet. 7:e1002075.

Roth, F. P., H. D. Lipshitz, and B. J. Andrews. 2009. Q&A: epistasis. J. Biol. 8:35.

Salverda, M. M., E. Dellus, F. A. Gorter, A. J. M. debets, J. van der Oost, R. F. Hoekstra, D. S. Tawfik, and J. A. G. M. de Visser. 2011. Initial mutations direct alternative pathways of protein evolution. PLoS Genet. 7:e1001321.

Sanjuán, R., A. Moya, and S. F. Elena. 2004. The contribution of epistasis to the architecture of fitness in an RNA virus. Proc. Natl. Acad. Sci. 101:15376–15379.

Segrè, D., A. DeLuna, G. M. Church, and R. Kishony. 2005. Modular epistasis in yeast metabolism. Nat. Genet. 37:77–83.

Sella, G. 2009. An exact steady state solution of Fisher's geometric model and other models. Theor. Popul. Biol. 75:30–34.

Sellis, D., B. J. Callahan, D. A. Petrov, and P. W. Messer. 2011. Heterozygote advantage as a natural consequence of adaptation in diploids. Proc.Natl. Acad. Sci. 108:20666–20671.

Sideraki, V., W. Huang, T. Palzkill, and H. F. Gilbert. 2001. A secondary drug resistance mutation of TEM-1 β-lactamase that suppresses misfolding and aggregation. Proc. Natl. Acad. Sci. USA 98:283–288.

Sokal, R. R., and F. J. Rohlf. 1995. Biometry. W.H. Freeman and Company, New York.

Szathmáry, E. 1993. Do deleterious mutations act synergistically? Metabolic control theory provides a partial answer. Genetics 133:127–132.

Tenaillon, O., O. K. Silander, J.-P. Uzan, and L. Chao. 2007. Quantifying organismal complexity using a population genetic approach. PLoS One 2:e217.

Tokuriki, N., F. Stricher, L. Serrano, and D. S. Tawfik. 2008. How protein stability and new functions trade off. PLoS Comp. Biol. 4:e1000002.

Tomatis, P. E., R. M. Rasia, L. Segovia, and A. J. Vila. 2005. Mimicking natural selection in metallo-β-lactamases through second-shell ligand mutations. Proc. Natl. Acad. Sci. USA 102:13761–13766.

Trindade, S., A. Sousa, K. B. Xavier, F. Dionisio, M. G. Ferreira, and I. Gordo. 2009a. Data from: positive epistasis drives the acquisition of multidrug resistance. doi:10.5061/dryad.20k7j.

Trindade, S., A. Sousa, K. B. Xavier, F. Dionisio, M. G. Ferreira, and I. Gordo. 2009b. Positive epistasis drives the acquisition of multidrug resistance. PLoS Genet. 5:e1000578.

Wang, Z., G. Minasov, and B. K. Shoichet. 2002. Evolution of an antibiotic resistance enzyme constrained by stability and activity trade-offs. J. Mol. Biol. 320:85–95.

Waxman, D. 2006. Fisher's geometric model of evolutionary adaptation—beyond spherical geometry. J. Theor. Biol. 241:887–895.

Waxman, D., and J. J. Welch. 2005. Fisher's microscope and Haldane's ellipse. Am. Nat. 166:447–457.

Weinreich, D. M. 2010. Predicting molecular evolutionary trajectories in principle and in practice. Encyclopedia of life sciences (ELS). John Wiley & Sons, Chichester, U.K.

Weinreich, D. M., N. F. Delaney, M. A. DePristo, and D. L. Hartl. 2006. Darwinian evolution can follow only very few mutational paths to fitter proteins. Science 312:111–114.

Weinreich, D. M., R. A. Watson, and L. Chao. 2005. Sign epistasis and genetic constraint on evolutionary trajectories. Evolution 59:1165–1174.

Welch, J. J., and D. Waxman. 2003. Modularity and the cost of complexity. Evolution 57:1723–1734.

Yokoyama, S., H. Yang, and W. T. Starmer. 2008. Molecular basis of spectral tuning in the red- and green-sensitive (M/LWS) pigments in vertebrates. Genetics 179:2037–2043.

Zeldovich, K. B., P. Q. Chen, and E. I. Shakhnovich.2007. Protein stability imposes limits on organism complexity and speed of molecular evolution. Proc. Natl. Acad. Sci. 104:16152–16157.

Zhou, H.-Z., S. T. Wlodek, and J. A. McCammon. 1998. Conformational gating as a mechanism for enzyme specificity. Proc. Natl. Acad. Sci. USA 95:9280–9283.

Associate Editor: J. Kelly