CrossMark

The Influence of Higher-Order Epistasis on Biological Fitness Landscape Topography

Daniel M. Weinreich[1,2] · **Yinghong Lan[1]** ·
Jacob Jaffe[1] · **Robert B. Heckendorn[3]**

Abstract The effect of a mutation on the organism often depends on what other mutations are already present in its genome. Geneticists refer to such mutational interactions as epistasis. Pairwise epistatic effects have been recognized for over a century, and their evolutionary implications have received theoretical attention for nearly as long. However, pairwise epistatic interactions themselves can vary with genomic background. This is called higher-order epistasis, and its consequences for evolution are much less well understood. Here, we assess the influence that higher-order epistasis has on the topography of 16 published, biological fitness landscapes. We find that on average, their effects on fitness landscape declines with order, and suggest that notable exceptions to this trend may deserve experimental scrutiny. We conclude by highlighting opportunities for further theoretical and experimental work dissecting the influence that epistasis of all orders has on fitness landscape topography and on the efficiency of evolution by natural selection.

Keywords Higher-order epistasis · Fitness landscapes topography · Natural selection · NK landscape · Sequence space combinatorics

✉ Daniel M. Weinreich
  Daniel_Weinreich@Brown.edu
  http://brown.edu/research/labs/weinreich/home

[1] Department of Ecology and Evolutionary Biology, Brown University, Providence, RI 02912, USA

[2] Center for Computational Molecular Biology, Brown University, Providence, RI 02912, USA

[3] Computer Science Department, University of Idaho, 875 Perimeter Drive, MS 1010, Moscow, ID 83844, USA

1 Introduction

One of the more evocative pictures of biological evolution is that of a population climbing the fitness landscape [37,44]. This image was originally proposed by Wright [73] to build intuition into his [72] and Fisher's [19] technical treatment of Darwin's theory of natural selection in finite populations under Mendelian genetics [51]. The topography of the fitness landscape represents the strength and direction of natural selection as local gradients that influence the direction and speed with which populations evolve.

While several distinct framings of the fitness landscape have been suggested [51], here we employ the projection of genotypic fitness over Maynard Smith's sequence space [36]. Sequence space is a discrete, high-dimensional space in which genotypes differing by exactly one point mutation are spatially adjacent. Thus, proximity on the fitness landscape corresponds to mutational accessibility, and selection will try to drive populations along the locally steepest mutational trajectory. (See [68] for several processes not readily captured by this construction.)

The most obviously interesting topographic feature of the fitness landscape is the number of maxima, a point already recognized by Wright [73]. Two (or more) maxima can constrain natural selection's ability to discover highest-fitness solutions, since populations may be required to transit lower-fitness valleys on the landscape en route. (Though see [25,65] for the population genetics of that process, sometimes called stochastic tunneling [15,25].)

1.1 Epistasis and Fitness Landscape Topography

Epistasis is the geneticist's term for interactions among mutational effects on the organism [46]. For example, genetically disabling two genes whose products act in the same linear biochemical pathway can have a much more modest effect than the sum of the effects of disabling either gene in isolation. Alternatively, disabling two functionally redundant genes can have a much more substantial effect than expected. (Indeed, such observations have taught us quite a bit about the organization of biochemical pathways, e.g., [2].)

Epistatic interactions between mutations can occur for any organismal trait, including fitness. Importantly, epistasis for fitness has an intimate connection to the topography of the fitness landscape, a fact also already appreciated by Wright [73]. For example, multiple peaks require the presence of mutations that are only conditionally beneficial (called sign epistasis [49,68]). More generally, an isomorphism exists between fitness landscapes defined by mutations at some L positions in the genome and the suite of epistatic interactions possible among them. This follows because, while any particular mutation can appear on 2^{L-1} different genetic backgrounds (assuming two alternative genetic states, or alleles, at each position), each such mutation-by-background pair corresponds to a distinct adjacency in sequence space. Consequently, arbitrary differences in the fitness effect of a mutation across genetic backgrounds can generically be represented on the fitness landscape [68].

1.2 Higher Order Epistasis and Fitness Landscape Topography

Widespread epistasis between pairs of mutations has been recognized in nature for over 100 years [46,67], and the corresponding evolutionary theory is fairly advanced (e.g., [5,71]). However, pairwise interactions can themselves vary with genetic background, called higher-order epistasis [13,67]. And while it is now becoming clear that higher-order interactions are commonplace in nature [32,42,61,67], their influence on natural selection is less well

understood (though see [55]). Here, we present a simple framework for assessing the influence on fitness landscape topography of epistatic terms of arbitrary order. We speculate that epistatic influence on the topography of naturally occurring fitness landscapes should decline with epistatic order. We tested this prediction using 16 published biological fitness landscapes.

2 Methods

2.1 The Order of Epistatic Interactions

Any set of L biallelic loci defines 2^L genotypes, each with 2^L potentially independent fitness values. Simultaneously, there are $\binom{L}{k}$ distinct subsets of k mutations that in principle can also independently contribute to a genotype's fitness. In total, there are thus $\sum_{k=0}^{L} \binom{L}{k}$ $= 2^L$ subsets of mutations (i.e., the power set of L mutations). This counting reflects the isomorphism between any fitness landscape and its corresponding suite of epistatic terms [67].

We designate interactions among any subset of k mutations as kth-order epistasis. Note that here first-order "epistasis" is degenerate in the sense that it represents the fitness effects of each of the L mutations in isolation. And our zeroth-order "epistatic" term is the benchmark, relative to which the effect of each subset of mutations is computed.

2.2 The Fourier–Walsh Transformation

Following earlier work [22,41,59,64,67] we employ the Fourier–Walsh transformation (Fig. 1a) to convert between fitness landscapes and their corresponding epistatic terms. This is a linear transformation written

$$\overrightarrow{E_W} = \frac{1}{2^L} \mathbf{\Psi} \overrightarrow{W}. \tag{1}$$

Here $\overrightarrow{W}$ is the vector of all 2^L fitness values arranged in the canonical order defined by ascending L-bit binary numbers encoding the corresponding genotype with respect to the presence or absence of each mutation (e.g., [33]). (W is the traditional population genetics symbol for fitness.) $\mathbf{\Psi}$ is the Hadamard matrix, the unique, symmetric $2^L \times 2^L$ matrix whose entries are either $+1$ or -1 and whose rows (and columns) are mutually orthogonal. ($\mathbf{\Psi}$ can be written for arbitrary L, as for example with the hadamard() function in the software package Matlab, Mathworks, Natick, MA.) Finally, $\overrightarrow{E_W}$ is the resulting vector of 2^L epistatic terms arranged in the canonical order defined by ascending L-bit binary numbers whose 1's indicate the corresponding subset of interacting loci. Figure 1a illustrates this transformation using the data in [45]. For example, the fourth component of $\overrightarrow{E_W}$ (-0.1429) signals a negative epistatic interaction between the two most 3' mutations in that dataset. (See Fig. 1 in [54] for a graphical representation of the elements of $\overrightarrow{E_W}$, and [50] for the relationship between Eq. (1) and other formalisms for computing epistatic terms.)

The orthogonality and symmetry of $\mathbf{\Psi}$ means that $\mathbf{\Psi}^T \cdot \mathbf{\Psi} = \mathbf{\Psi}^2 = 2^L \mathbf{I}$, where $\mathbf{I}$ is the identity matrix. This means that, just as Eq. (1) converts any landscape into its epistatic

a Compute epistatic coefficients

| $\frac{1}{2^L}$	×	Ψ	×	$\vec{W}$	=	$\vec{E}_w$	Binary String Encoding	Epistatic Order
$\frac{1}{64}$	×	$\begin{bmatrix} +1 & +1 & +1 & +1 & +1 & +1 & +1 & \dots & +1 \\ +1 & -1 & +1 & -1 & +1 & -1 & +1 & \dots & -1 \\ +1 & +1 & -1 & -1 & +1 & +1 & -1 & \dots & -1 \\ +1 & -1 & -1 & +1 & +1 & -1 & -1 & \dots & +1 \\ +1 & +1 & +1 & +1 & -1 & -1 & -1 & \dots & -1 \\ +1 & -1 & +1 & -1 & -1 & +1 & +1 & \dots & -1 \\ +1 & +1 & -1 & -1 & -1 & -1 & +1 & \dots & +1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ +1 & -1 & -1 & +1 & -1 & +1 & +1 & \dots & -1 \end{bmatrix}$	×	$\begin{bmatrix} 1.0137 \\ 2.5027 \\ 1.2963 \\ 2.9960 \\ 2.8863 \\ -2 \\ 1.3387 \\ \vdots \\ 2.7427 \end{bmatrix}$	=	$\begin{bmatrix} 2.2880 \\ 0.0857 \\ -0.1007 \\ -0.1429 \\ 0.0201 \\ -0.4814 \\ -0.2020 \\ \vdots \\ -0.1051 \end{bmatrix}$	$\begin{bmatrix} 000000 \\ 000001 \\ 000010 \\ 000011 \\ 000100 \\ 000101 \\ 000110 \\ \vdots \\ 111111 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 1 \\ 1 \\ 2 \\ 1 \\ 2 \\ 2 \\ \vdots \\ 6 \end{bmatrix}$

b Compute epistatic contributions to landscape topography

1. Record epistatic order of each element in $\vec{E}_w$ after sorting from largest to smallest in absolute value.
2. For each $2 \leq m \leq 2^L$
   a. Construct $\vec{W}^{(m)}_{best} = \Psi \, \vec{E}^{(m)}_w$ using the m largest elements in $\vec{E}_w$.
   b. Record residual variance between $\vec{W}^{(m)}_{best}$ and $\vec{W}$.
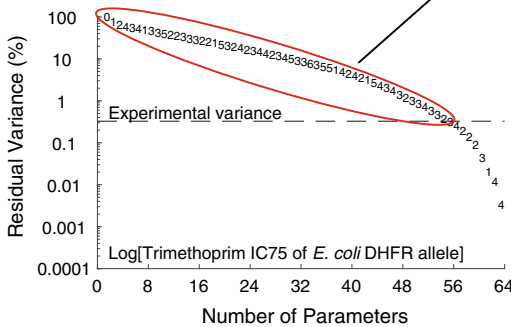
c Compute correlation coefficient τ_b

$\begin{bmatrix} 0 \\ 1 \\ 2 \\ 4 \\ 3 \\ 4 \\ 1 \\ \vdots \\ 3 \\ 2 \end{bmatrix}$ $\begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ \vdots \\ 4 \\ 4 \end{bmatrix}$

$\tau_b(\text{Obs,Exp}) = 0.1921$

$P = 0.03639$

Observed Expected
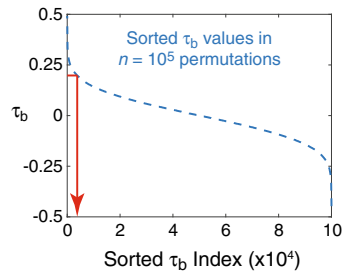
Fig. 1 Analytic pipeline, illustrated with data from Palmer et al. [45]. **a** For each dataset, published fitness data (or a suitable proxy, written $\vec{W}$) were first converted to the corresponding epistatic terms ($\vec{E}$ using the Fourier–Walsh transformation (Eq. 1). **b** Explanatory power of a succession of models using only the m largest epistatic terms in absolute value ($\overrightarrow{W^{(m)}_{best}}$) were compared with the published data. For given value of m, these models provably have the greatest explanatory power (smallest residual variance) of any model with exactly m parameters (Appendix). The symbols plotted represent the epistatic order (Sect. 1.2) of each successive parameter added to the model. **c** Rank correlation coefficient (τ_b) between the empirical sequence of epistatic orders and those of our naïve expectation (Eq. 2) were computed. In cases where experimental variance was reported, these sequences were truncated as soon as the remaining model variance was less than the experimental variance. For the data shown, that truncation occurred after the 55th epistatic term. Finally, statistical significance was assessed by a permutation test that asked whether the observed sequence of epistatic orders was significantly different than random. For the data shown (red arrow), the observed value of τ_b (0.1921) was smaller than only the 3639 largest of 10^5 values obtained by the permutation test, yielding $P = 0.03639$

terms, so too can any vector of epistatic terms $\vec{E}$ be converted into its corresponding fitness landscape as $\overrightarrow{W} = \Psi \vec{E}$. We take advantage of this fact next.

2.3 Subsetting Approximations of a Fitness Landscape

Given fitness function $\overrightarrow{W}$, we now introduce subsetting approximations $\overrightarrow{W^{(m)}} = \Psi \overrightarrow{E_W^{(m)}}$. Here, the $\overrightarrow{E_W^{(m)}}$ are constructed so that $0 \leq m \leq 2^L$ of the components are from $\overrightarrow{E_W} = \frac{1}{2^L} \Psi \overrightarrow{W}$ (Eq. 1) and the remaining $2^L - m$ components are set to zero. There are thus 2^{2^L} subsetting approximations for any fitness function $\overrightarrow{W}$ (corresponding to the power set of the 2^L epistatic terms in $\overrightarrow{E_W}$). As a consequence of the orthogonality of the Fourier–Walsh transformation, the sum of squares distance between fitness function $\overrightarrow{W}$ and subsetting approximation $\overrightarrow{W^{(m)}} = \Psi \overrightarrow{E_W^{(m)}}$ is minimized for given m if and only if $\overrightarrow{E_W^{(m)}}$ uses the m largest components in absolute value of $\overrightarrow{E_W}$ (see Appendix). We denote these $0 \leq m \leq 2^L$ best subsetting approximations $\overrightarrow{W_{\text{best}}^{(m)}}$.

(Subsetting approximations defined by interaction order rather than absolute magnitude of epistatic terms were recently employed elsewhere [55].)

2.4 Quantifying the Influence of Epistatic Terms on Empirical Fitness Landscape Topography

To examine the influence of epistasis on fitness landscape topography as a function of epistatic order, we first used Eq. (1) to compute $\overrightarrow{E_W}$ for each $\overrightarrow{W}$ gleaned from the literature (Sect. 2.7). For each $1 \leq m \leq 2^L$, we then iteratively constructed each $\overrightarrow{W_{\text{best}}^{(m)}}$. Finally, for each m we recorded the residual variance between $\overrightarrow{W_{\text{best}}^{(m)}}$) and $\overrightarrow{W}$ (minimized by this subsetting approximation; Sect. 2.3), together with the epistatic order of the mth-largest component of $\overrightarrow{E_W}$. Figure 1b illustrates this process.

2.5 Statistics

Our hypothesis is that the influence of an epistatic term on the fitness landscape should decline with epistatic order. Put another way, we expected that after sorting the elements of $\overrightarrow{E_W}$ (Eq. 1) by their absolute magnitudes, the associated epistatic orders should be represented by a vector of 2^L integers that reads:

$$0, 1, 1, \ldots, 2, 2, 2, \ldots, \ldots L - -1, L - -1, \ldots, L. \tag{2}$$

Specifically, this vector consists of one zero, followed by L ones, $\binom{L}{2}$ twos and in general $\binom{L}{k}$ k's for all $0 \leq k \leq 2^L$.

We tested this hypothesis for each dataset by first computing Kendall's τ_b correlation coefficient [28] between this expectation and the epistatic orders observed among the elements in $\overrightarrow{E_W}$ sorted by absolute magnitude. τ_b is one (negative one) when the observed epistatic orders are perfectly correlated (anticorrelated) with expectation, and zero when they are uncorrelated. Note that Kendall's τ_b statistic is appropriate because it accommodates ties.

For studies that also reported experimental variance, we computed the correlation coefficient after discarding the epistatic orders of all j elements in $\overrightarrow{E_W}$ that reduced residual variance by less than experimental variance (see Fig. 1b and Table 1) as well as the last j epistatic order values in our expectation (given by Eq. 2).

For each dataset, we then used a permutation test to test the null hypothesis that the corresponding correlation coefficient is zero. Specifically, each dataset is characterized by some number of epistatic terms: 2^L in cases where no experimental variance estimate is provided, or $2^L - j$ in cases where we were able to identify non-significant epistatic components (see previous paragraph and Table 1). For each of $n = 10^5$ replicates, we computed the rank correlation coefficient between two random permutations of this number (2^L or $2^L - j$) of the epistatic order values drawn from Eq. (2) for given L. We then sorted correlation coefficients, and the uncorrected P value reported for each dataset (Table 1) was taken as the fraction of permutations in which a correlation coefficient greater than or equal to the empirical value was observed. Thus, ours is a one-tailed test of the hypothesis that no positive correlation is present. This process is illustrated in Fig. 1c.

We used the Bonferroni–Holm method [24] to correct for multiple tests. In addition, under the null hypothesis that epistatic orders are uncorrelated with the naïve expectation given by Eq. (2), the distribution of P values observed across datasets should be uniformly distributed. We tested this hypothesis with a G-test after binning counts of empirically observed P values. We assessed statistical significance relative to the χ^2 distribution [56].

2.6 Empirical Datasets

To compute all 2^L epistatic terms in a fitness landscape defined over L biallelic loci requires data on the fitness values (or suitable proxy) for each of the corresponding 2^L genotypes. We previously designated such datasets combinatorially complete [67], and the datasets analyzed here are shown in Table 1. Several datasets [4,34,43,45] had a few loci with cardinality greater than two. In these cases, we examined one "slice" through the landscape defined by randomly choosing just two alleles at those loci.

Several studies examined multiple phenotypes for a single set of mutations, and follow-up studies sometimes presented additional phenotypes for a previously described set of mutations. Those cases are enumerated in Table 2; for each set of mutations we randomly sampled just one phenotype. Table 2 also lists all combinatorially complete datasets we know that are defined over loci with cardinality greater than two. These were excluded here because the Fourier–Walsh framework doesn't trivially generalize to higher cardinalities.

Following [67], datasets reporting growth rates [4,10,14,16,20,21,70] and drug-resistance phenotypes [8,33,38,39,45,66] were log-transformed before analysis. Following [45], negative two was used in place of log-transformed values when growth rate or drug resistance phenotypes of zero were observed. (In all cases, this is roughly one log order smaller than the smallest non-zero log-transformed value.) In cases where only mean and experimental variances (but not individual replicate observations) were provided, log transformations were approximated by Taylor expansions: $\overline{\ln(x)} \approx \ln(\bar{x}) - s_x^2/2\bar{x}^2$ and $s_{\ln(x)}^2 \approx (s_x/\bar{x})^2$. In cases where only means (but not variances) were provided, log transformations were approximated as $\overline{\ln(x)} \approx \ln(\bar{x})$.

Following [45], for studies in which experimental variance estimates were provided, we recorded this quantity as a fraction of the total model variance. In one case [8], standard error was reported as standard error over "at least" two replicates; we therefore assumed $n = 2$ for each observation in that dataset. In one case [29], 95% experimental confidence intervals were

Table 1 Analyses of published combinatorially complete empirical and simulated (NK) fitness landscapes, sorted by P value associated with Kendall's τ_b

Phenotype [citation]	Number of loci (L)	Number of maxima	Number of epistatic terms significantly different from zero[a]	Kendall's τ_b correlation coefficient	P value[b]
Log[S. cerevisiae HSP90 mutant growth rates] [4]	6	4	41	0.6202	< 0.00001***
Log[diploid S. Cerevisiae mutant growth rate] [21]	6	4	3	0.6667	<0.00001***
Log[E. coli IMDH mutant relative growth rates] [34]	6	1	N.D.[c]	0.7566	< 0.00001***
Avian lysozyme thermostability [35]	3	1	4	0.5	<0.00001***
$N = 5$, $K = 0$	5	1	N.D.[d]	0.3333	<0.00001***
Log[relative fitness among Methylobacterium extorquens mutants] [10]	4	1	N.D.[c]	0.7527	0.00001***
Log[HIV replicative capacity on CCR5+ cells] [14]	5	3	25	0.5703	0.00002***
Log[cefotaxime MIC of E. coli TEM alleles] [66]	5	1	29	0.5490	0.00011**
Log[relative viability among fruit fly mutants] [70]	5	3	N.D.[c]	0.4896	0.00027**
Log[cefalexin MIC of Bacillus cereus metallo-β-lactamase alleles] [38]	4	1	N.D.[c]	0.5376	0.00448
$N = 5$, $K = 1$	5	2	N.D.[d]	0.5714	0.00474
Log[relative fitness among LTEE E. coli mutants in DM25 + EGTA] [20]	5	2	30	0.3486	0.01023
$N = 5$, $K = 2$	5	2	N.D.[d]	0.4428	0.0156

Table 1 continued

Phenotype [citation]	Number of loci (L)	Number of maxima	Number of epistatic terms significantly different from zero[a]	Kendall's τ_b correlation coefficient	P value[b]
Log[relative colony growth rate among *Aspergillus niger* mutants] [16]	5	4	10	0.4387	0. 02002
Percent production of 5-epi-aristolochene by sesquiterpene synthase mutants [43]	6	10	N.D.[c]	0.1974	0.02639
Log[pyrimethamine IC50 of *P. falciparum* DHFR alleles assayed in *E. coli*] [33]	4	2	14	0.4337	0.03133
Log[IC75 of *E. coli* DHFR alleles against trimethoprim] [45]	6	2	55	0.1921	0.03639
$N = 5$, $K = 4$	5	5	N.D.[d]	0.1735	0.1442
Mammalian glucocorticoid receptor cortisol sensitivity [7]	4	4	N.D.[c]	0.1075	0.30182
Log[MIC of *E. coli* TEM alleles against ampicillin] [39]	4	3	N.D.[c]	0.0430	0.41356
$N = 5$, $K = 5$	5	7	N.D.[d]	−0.1632	0.86161

[a]Maximum possible value is 2^L

[b]Uncorrected value from permutation test ($n = 10^5$ replicates). Bonferroni-corrected P values: *** ≤ 0.001; $0.001 < ** \leq 0.01$; $0.01 < * \leq 0.05$

[c]No data because no experimental variance estimates provided with this dataset

[d]No data because simulated fitness landscapes have no experimental variance

Table 2 Published combinatorially complete fitness landscapes not examined here

Phenotype [citation]	Number of loci	Number of genotypes[a]
Cycloguanil IC50 of *S. cerevisiciae* carrying *P. falciparum* DHFR alleles[b] [11]	3	$2^3 = 8$
Pyrimethamine IC50 of *S. cerevisiciae* carrying *P. falciparum* DHFR alleles[b] [8]	3	$2^3 = 8$
MIC against pyrimethamine of *S. cerevisiae* carrying *P. falciparum* DHFR alleles[b] [8]	3	$2^3 = 8$
Pyrimethamine IC50 of *P. vivax* DHFR alleles assayed in *S. cerevisiae*[b] [26]	4	$2^4 = 16$
MIC of TEM β-lactamase mutants against 15 β-lactams[b] [39]	4	$2^4 = 16^c$
E. coli operator binding affinities for 400 repressor sequences [48]	4	$4^4 = 256$
Relative fitness among LTEE *E. coli* mutants in DM25[b] [29]	5	$2^5 = 32$
Relative fitness among LTEE *E. coli* mutants in DM25 + guanazole[b] [20]	5	$2^5 = 32$
HIV replicative capacity on CXCR5+ cells[b] [14]	5	$2^5 = 32$
Transcription factor/response element specificity in an ancient steroid hormone receptor [1]	5	$4^2 \times 2^3 = 128^d$
MIC of TEM-β-lactamase mutants against Piperacillin supplemented with clavulanic acid[b] [63]	6	$2^5 = 32$
Diploid *S. cerevisiae* mutant growth rate[b] [21]	6	$2^6 = 64$
Percent production of minor products by sesquiterpene synthase mutants[b] [43]	6	$2^6 = 64$
Percent production of 4-EE by sesquiterpene synthase mutants[b] [43]	6	$2^6 = 64$
Percent production of PSD by sesquiterpene synthase mutants[b] [43]	6	$2^6 = 64$
104 mouse DNA-binding proteins' affinity for 10 bp binding motifs [3]	10	$4^{10} = 1{,}048{,}576^e$
GFP affinity for 10 nucleotide base pair binding motifs [52]	10	$4^{10} = 1{,}048{,}576$
Affinity of 1032 DNA-binding proteins spanning eukaryotic diversity against 10 nucleotide base pair binding motifs [69]	10	$4^{10} = 1{,}048{,}576^f$

[a]Written as the product of cardinalities across loci
[b]Another phenotype from this system is included in Table 1
[c]In total, 16×15 distinct β-lactam compounds = 240 observations are reported in this study
[d]This study examined all combinations of 4 nucleotides at two key positions in the DNA response element together with all combinations of two amino acids at three key positions in the transcription factor
[e]In total $1{,}048{,}576 \times 104$ DNA-binding proteins = 109,051,904 observations are reported in this study
[f]In total $1{,}048{,}576 \times 1{,}032$ DNA-binding proteins = 1,082,130,432 observations are reported in this study

reported, so variance estimates were computed under the assumption of normally distributed noise as $s^2 = (n \cdot CI95/1.96)^2$.

2.7 Simulated Fitness Landscapes

We used *NK* fitness landscapes [27] to test our hypothesis in a framework with explicitly tunable mutational interactions. Genomes in the *NK* model carry *N* loci. The fitness contribution of each locus depends on its allelic state and that at *K* others. Thus $0 \leq K \leq N - 1$ represents a parameter that tunes the level of epistatic interaction in the landscape. (See [41] and references therein for a number of elegant statistical properties of *NK* fitness landscapes.) We set $N = 5$ and generated one *NK* landscape for each $0 \leq K \leq N - 1$, where interacting loci were assigned at random in the genome. Simulated data were then analyzed as described in Fig. 1.

2.8 Data and Software Archiving

Input data files, together with purpose-built MatLab code to perform all analyses described are archived at https://github.com/weinreichlab/JStatPhys2018. Kendall's τ_b correlation coefficient was computed using MatLab code developed elsewhere [9]. *NK* fitness landscapes were generated using code downloaded from https://github.com/qzcwx/NK-generator.

3 Results

Epistasis can have profound consequences at many levels of biological organization [47, 53, 60, 71]. Here we tested the hypothesis that the influence of epistasis on empirical fitness landscape topography should decline as a function of epistatic order.

This study was originally stimulated by Fig. 2 in Palmer et al. [45], which examined six mutations in the dihydrofolate reductase (DHFR) gene of *E. coli* that contribute to increased resistance to an antimicrobial called trimethoprim. In that analysis, particular second- and third-order interactions were the third- and second-most influential epistatic terms for fitness landscape topography respectively. Indeed, just two of the first ten most influential epistatic terms were first-order, and in aggregate first-order terms explained just $\sim 28\%$ of the variance in fitness across the landscape. At first blush, these results seem to challenge the hypothesis outlined in the previous paragraph, and we therefore sought to explore the pattern more broadly using published data from other systems.

Fig. 2 Distribution of uncorrected *P* values among 16 empirical datasets. Under any null model, *P* values are expected to be uniformly distributed (black bars; note both axes are log-transformed). Instead observed *P* values (grey bars) are sharply skewed toward small values ($G = 143.77$, $P_{\text{d.f.}=5} \ll 0.01$)
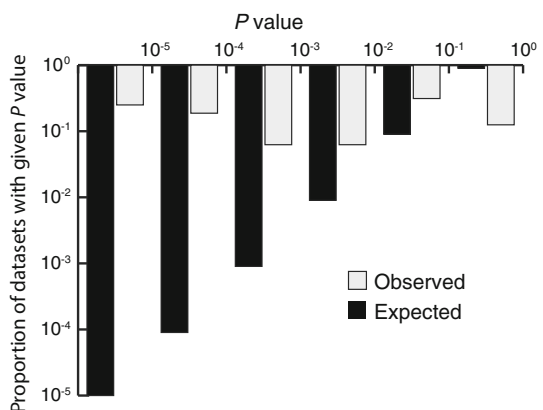
Figure 1 illustrates the application of our analytic pipeline (see Sect. 2) to these same data. Our Fig. 1b closely recapitulates Fig. 2a in Palmer et al. [45]. While the precise sequence of epistatic terms differs slightly (likely because the previous study employed a subtly different framework for computing epistatic terms), higher-order epistatic interactions are again responsible for some the largest reductions in residual variance. Indeed, as previously observed, just two of the first ten terms are first-order, and in aggregate and first-order terms again explain just $\sim 28\%$ of the variance in the data (Table 3a, compare the first two columns with Fig. 2b in [45]). Importantly however, Fig. 1c illustrates that we find a significant, positive correlation between expectation (Eq. 2) and the observed influence of epistatic terms on landscape topography as a function of their order ($\tau_b = 0.1921$, $P = 0.03639$).

We next applied our pipeline to 15 other published, combinatorially complete datasets. Results are summarized in Table 1 and shown graphically in Fig. S1. Out of all 16 datasets examined, 14 exhibit a significantly positive correlation between observation and the expectation, and eight of these remain significant after Bonferroni correction for multiple tests. Moreover, across datasets Table 1 exhibits a bias toward small P values. Under the null hypothesis (no significant correlation with expectation), we would expect a uniform distribution of P values. Instead, the observed distribution is sharply and significantly skewed toward small values (Fig. 2, $G = 143.77$, $P_{d.f.=5} \ll 0.01$).

We also applied our pipeline to *NK* fitness landscapes generated for $N = 5$ and $0 \le K \le N - 1$. We set $N = 5$ because the average size of the empirical datasets is 4.875 loci. Those results are also included in Table 1 (though omitted from Fig. 2).

4 Discussion

Using a novel analytic pipeline (Fig. 1), we have examined 16 published, combinatorially complete biological datasets. This analysis broadly confirms our intuition that the influence of epistatic terms on empirical fitness landscape topography should decline with order, i.e., with the number of interacting mutations. Consistent with this intuition, observed fit to expectation in our simulated (*NK*) fitness landscapes deteriorates as the amount of epistasis (K) goes up. In the limit of $K = N - 1$, fitness values (and hence, our epistatic terms) are i.i.d., and consequently the correlation is ~ 0.

While considerable heterogeneity in effect exists among our empirical datasets (Table 1), eight of the 16 exhibit a Bonferroni-corrected, significantly positive correlation with expectation (Eq. 2). Moreover, across all 16 empirical datasets, we find a sharp bias toward significant P values (Fig. 2). Nor is there any correlation between the size of the dataset and uncorrected P value (not shown), suggesting that low statistical power is unlikely to contribute to the overall picture.

The relative magnitudes of epistatic terms depend on the underlying fitness scale employed [30,67]. Although we log-transformed growth rate and drug resistance data (see Sect. 2.6), we have otherwise overlooked this fact. Recently, approaches for systematically rescaling data to minimize higher-order epistatic effects have been introduced [54] (see also [41,62]). Applications of such methods would certainly have quantitative consequences for results presented here. However, because these approaches (on average) reduce higher-order epistatic terms, we believe this omission renders our conclusions conservative.

We also acknowledge that we failed to honor experimental uncertainty in the magnitudes of epistatic effects observed, which would almost certainly weaken the signal reported in

Table 1. While we regard a rigorous treatment of experimental noise to be outside the scope of the present study, we note that the results presented in Fig. 2 are robust to its influence. Nevertheless, this is a serious concern for future consideration: because epistasis represents the difference between mutational fitness effects on different genetic background, experimental variance in fitness assays must be summed when computing variance in epistatic terms. For example, variance in epistatic terms computed with Eq. (1) will be roughly 2^L as large as variance in the individual, underlying fitness measurements. Recently, an alternative, ranks-based approach to assessing epistatic interactions between mutations has been proposed [13], which appears to be less sensitive to this effect.

4.1 The Combinatorics of Higher-Order Epistasis

This work was originally stimulated by a previous study [45] that examined six mutations in the DHFR gene responsible for increased trimethoprim resistance in *E. coli*. At first blush, results summarized in Fig. 2 of that study called into question the hypothesis that higher-

Table 3 Average epistatic influence on fitness landscape topography as a function of epistatic order in select datasets

Epistatic order	Aggregate reduction in residual variance	Number of epistatic terms significantly different from zero	Mean reduction in residual variance per epistatic term[a]
(a)	Log[IC75 of *E. coli* DHFR alleles against trimethoprim] [45]		
First	0.279	5	**0.056**
Second	0.266	12	0.022
Third	0.233	18	0.013
Fourth	0.144	11	0.013
Fifth	0.0685	6	0.011
Sixth	0.0065	1	0.0065
(b)	Mammalian glucocorticoid receptor cortisol sensitivity [7]		
First	0.171	4	0.043
Second	.405	6	0.067
Third	0.420	4	**0.105**
Fourth	0.004	1	0.004
(c)	Log[MIC of *E. coli* TEM alleles against ampicillin] [39]		
First	0.353	4	0.088
Second	0.278	6	0.043
Third	0.279	4	0.070
Fourth	0.091	1	**0.091**
(d)	*N* = 5, *K* = 4		
First	0.027	5	0.005
Second	0.315	10	0.315
Third	0.402	10	0.402
Fourth	0.241	5	**0.412**
Fifth	0.015	1	0.015

[a]Largest value for each dataset shown in bold

order epistasis should only modestly influence naturally occurring fitness landscapes. And the salient features of that figure were recapitulated by our treatment (Fig. 1b, Table 3a).

However, our statistical analysis of those data reveals a strong positive correlation between epistatic influence on fitness topography as a function of epistatic order, consistent with our hypothesis (Fig. 1c). Thus in this system, the substantial influence of a few high-order epistatic terms is nevertheless consistent with the idea that high-order epistatic terms should in general only modestly contribute to fitness topography.

The resolution to this puzzle resides in the combinatoric number of epistatic terms. As noted above, given L biallelic loci there are $\binom{L}{k}$ epistatic coefficients of order k, and this quantity grows almost exponentially for $k \ll L$. Indeed, after normalizing the summed influence of all epistatic terms of order k by the number of such terms, we observe that the per-term effect declines almost monotonically with order in this dataset (Table 3a; see also [67]). More generally, in all but three of the datasets examined, the normalized explanatory power is largest for first-order epistatic terms. Intriguingly, those three exceptions (see Table 3b–d: mammalian glucocorticoid receptor cortisol sensitivity [7], log[MIC of *E. coli* TEM allele sensitivity to ampicillin] [39] and the $N = 5$, $K = 4$ simulated fitness landscape) correspond to the three datasets with the largest P values in Table 1.

The consideration of the combinatorics of the problem is closely related to the Fourier spectrum of a fitness landscape [41,57], namely the sum of squared epistatic coefficients as a function of interaction order. (This connection derives formally from the Appendix, which implies that the squared magnitude of each epistatic coefficient is monotonic in its influence on landscape topography.) The Fourier spectrum is proportional to the binomial coefficient when each genotype's fitness is identically and independently distributed. This follows from the fact that on such landscapes all epistatic coefficients are also i.i.d., together with the combinatorics outlined in the previous paragraph. But as already anticipated by results in Table 3, the Fourier spectrum for the DHFR datasets is sharply shifted toward lower-order terms (not shown), as has previously been reported for both sesquiterpene synthase and several others biological datasets [41].

Nevertheless, declining average epistatic effects notwithstanding we find many examples of specific epistatic terms with anomalously large explanatory effects in many of the datasets examined here (Fig. S1). We suggest that these may reflect important mechanistic interactions among those particular mutations in the underlying biology of the system, thus representing potentially fruitful entry points for the molecular biologist [17].

4.2 Epistasis and the Efficiency of Natural Selection

Our observation that the influence of epistatic terms on naturally occurring fitness landscapes declines with epistatic order raises the question of how epistatic terms influence the efficiency of natural selection. We lack a complete theoretical understanding of this connection.

One well-developed result concerns the influence of epistasis on the selective accessibility of mutational trajectories to high fitness genotypes. First, sign epistasis means that the sign of the fitness effect of a mutation varies with genetic background [68], and it renders selectively inaccessible at least some mutational trajectories to high fitness (e.g., [66]). But connections between sign epistasis and epistatic order are only now being developed [13]. Second, a subsetting approach similar to ours (Sect. 2.3) was recently used to examine the influence of epistatic interactions selectively accessible mutational trajectories to high fitness genotypes [55] in six of the datasets described here. Those authors found that higher-order terms indeed substantially alter the identity of selectively favored mutational trajectories to high-fitness

genotypes, as well as their probabilities of realization. Further and consistent with findings here, that study also noted that the absolute magnitude of epistatic terms had an even larger effect on realized mutational trajectories than did their interaction order.

Moreover, pairwise epistasis has long been understood to influence not just the selective accessibility of high fitness genotypes but also the pace at which natural selection both increases the frequency of beneficial mutations (e.g., [18]) and at which it purges deleterious mutations (e.g., [31]). This work is closely related to the role that genetic recombination can play in "unlocking" epistatically interacting mutations (e.g., [5,40]). However, to our knowledge the relationship between these effects and higher-order epistasis remains entirely unexplored.

In addition, we have only quantitatively examined the sequence of epistatic orders sorted by explanatory power (Fig. 1c). Thus, a great deal of information present in these data (e.g., the slopes in Figs. 1b and S1) remains to be examined. And of course, the number and size of available combinatorially complete datasets continues to grow, motivating further work in this regard. It seems reasonable to suppose that the development and testing of more nuanced theoretical predictions may be possible using data of the sort examined here.

Finally, we note that the Fourier–Walsh framework employed here depends on the availability of combinatorially complete datasets. But the experimental demands of this approach grow exponentially with the number of mutations examined. This fact sharply limits the scalability of analytic pipelines like ours. Recently, theoretical progress has been made in the analysis of less-than-complete datasets [6,12,13], and older work has also explored this idea [23,58]. Theory that allows inferences using sparse datasets is likely to be a key advance in our ability explore broad, evolutionarily fascinating questions such as those considered here.

Appendix: The Explanatory Power of Fourier–Walsh Coefficients is Monotonic in Their Absolute Magnitude

Assume two fitness functions defined over L biallelic loci are represented as column vectors $\overrightarrow{W}$ and $\vec{X}$ with Fourier–Walsh coefficients $\overrightarrow{E_W}$ and $\overrightarrow{E_X}$ computed with Eq. (1). Define the sum of squares distance between $\overrightarrow{W}$ and $\vec{X}$ as $\|\overrightarrow{W} - \vec{X}\| \equiv \sum_{i=1}^{2^L}(W_i - X_i)^2 = (\overrightarrow{W} - \vec{X})^{\mathrm{T}} \cdot \overrightarrow{W} - \vec{X}$, where w_i and x_i are the ith components of $\overrightarrow{W}$ and $\vec{X}$, respectively.

Theorem 1 *Sum of squares distance equivalence*

$$\|\overrightarrow{W} - \vec{X}\| = 2^L \|\overrightarrow{E_W} - \overrightarrow{E_X}\|.$$

Proof By definition

$$\vec{W} = \mathbf{\Psi}\overrightarrow{E_W} \text{ and } \vec{X} = \mathbf{\Psi}\overrightarrow{E_X}$$

where $\mathbf{\Psi}$ is the Hadamard matrix (see Sect. 2.2).

Therefore

$$
\begin{aligned}
\|\vec{W} - \vec{X}\| &= \|\mathbf{\Psi}\overrightarrow{E_W} - \mathbf{\Psi}\overrightarrow{E_X}\| \\
&= \left(\mathbf{\Psi}\overrightarrow{E_W} - \mathbf{\Psi}\overrightarrow{E_X}\right)^{\mathrm{T}} \cdot \left(\mathbf{\Psi}\overrightarrow{E_W} - \mathbf{\Psi}\overrightarrow{E_X}\right) \\
&= \left(\left(\mathbf{\Psi}\overrightarrow{E_W}\right)^{\mathrm{T}} - \left(\mathbf{\Psi}\overrightarrow{E_X}\right)^{\mathrm{T}}\right) \cdot \left(\mathbf{\Psi}\overrightarrow{E_W} - \mathbf{\Psi}\overrightarrow{E_X}\right) \\
&= \left(\left(\overrightarrow{E_W}\right)^{\mathrm{T}}\mathbf{\Psi}^{\mathrm{T}} - \left(\overrightarrow{E_X}\right)^{\mathrm{T}}\mathbf{\Psi}^{\mathrm{T}}\right) \cdot \left(\mathbf{\Psi}\overrightarrow{E_W} - \mathbf{\Psi}\overrightarrow{E_X}\right) \\
&= \left(\left(\overrightarrow{E_W}\right)^{\mathrm{T}} - \left(\overrightarrow{E_X}\right)^{\mathrm{T}}\right)\mathbf{\Psi}^{\mathrm{T}} \cdot \mathbf{\Psi}\left(\overrightarrow{E_W} - \overrightarrow{E_X}\right)
\end{aligned}
$$

But recall that $\mathbf{\Psi^T\Psi} = 2^L\mathbf{I}$, where $\mathbf{I}$ is the identity matrix. Thus

$$
\begin{aligned}
\|\vec{W} - \vec{X}\| &= 2^L \left(\left(\overrightarrow{E_W}\right)^{\mathrm{T}} - \left(\overrightarrow{E_X}\right)^{\mathrm{T}}\right) \cdot \left(\overrightarrow{E_W} - \overrightarrow{E_X}\right) \\
&= 2^L \left(\overrightarrow{E_W} - \overrightarrow{E_X}\right)^{\mathrm{T}} \cdot \left(\overrightarrow{E_W} - \overrightarrow{E_X}\right) = 2^L\|\overrightarrow{E_W} - \overrightarrow{E_X}\|.
\end{aligned}
$$

$\square$

An interesting property of the Hadamard matrix is that $\mathbf{\Psi}^{\mathrm{T}} = 2^L\mathbf{\Psi}^{-1}$. Without the 2^L this equality is the hallmark of a rotational transformation. This means that Fourier–Walsh coefficients are simply the result of a high dimensional axis rotation of the coordinates of function space, together with a uniform contraction. This provides intuition into Theorem 1: rotating the space and contracting it uniformly only changes the distance between two vectors in the space by the constant of contraction.

Theorem 2 *Minimizing the sum-of-squares distance of subsetting approximations*

The subsetting approximation $\overrightarrow{W^{(m)}} = \mathbf{\Psi}\overrightarrow{E_W^{(m)}}$ that minimizes the sum of squares distance to function $\vec{W}$ is the one whose $\overrightarrow{E_W^{(m)}}$ uses the m largest components in absolute value in $\overrightarrow{E_W} = \frac{1}{2^L}\mathbf{\Psi}\vec{W}$.

Proof By Theorem 1, the sum of squares distance between $\vec{W}$ and $\overrightarrow{W^{(m)}}$ is $\|\vec{W} - \overrightarrow{W^{(m)}}\| = 2^L\|\overrightarrow{E_W} - \overrightarrow{E_W^{(m)}}\|$, which means that we can equivalently solve the minimization problem on either side of the equality. And trivially, the right-hand side is minimized when the m nonzero components in $\overrightarrow{E_W^{(m)}}$ are the m largest components in absolute value in $\overrightarrow{E_W}$. (The squaring of differences in epistatic terms in the definition of $\|\overrightarrow{E_W} - \overrightarrow{E_w^{(m)}}\|$ removes the significance of their sign.) $\square$

References

1. Anderson, D.W., McKeown, A.N., Thornton, J.W.: Intermolecular epistasis shaped the function and evolution of an ancient transcription factor and its DNA binding sites. eLife **4**, e07864 (2015)
2. Avery, L., Wasserman, S.: Ordering gene function: the interpretation of epistasis in regulatory hierarchies. Trends Genet. **8**, 312–316 (1992)
3. Badis, G., Berger, M.F., Philippakis, A.A., Talukder, S., Gehrke, A.R., Jaeger, S.A., Chan, E.T., Metzler, G., Vedenko, A., Chen, X., Kuznetsov, H., Wang, C.E., Coburn, D., Newburger, D.E., Morris, Q., Hughes, T.R., Bulyk, M.L.: Diversity and complexity in DNA recognition by transcription factors. Science **324**, 1720–1723 (2009)
4. Bank, C., Matuszewski, S., Hietpas, R.T., Jensen, J.D.: On the (un)predictability of a large intragenic fitness landscape. Proc. Natl. Acad. Sci. **113**, 14085–14090 (2016)
5. Barton, N.H.: Why sex and recombination? Cold Spring Harbor Symp. Quantit. Biol. **74**, 187–195 (2009)
6. Beerenwinkel, N., LPachter, L., Sturmfels, B.: Epistasis and shapes of fitness landscapes. Stat. Sin. **17**, 1317–1342 (2007)
7. Bridgham, J.T., Ortlund, E.A., Thornton, J.W.: An epistatic ratchet constrains the direction of glucocorticoid receptor evolution. Nature **461**, 515–519 (2009)
8. Brown, K.M., Costanzo, M.S., Xu, W., Roy, S., Lozovsky, E.R., Hartl, D.L.: Compensatory mutations restore fitness during the evolution of dihydrofolate reductase. Mol. Biol. Evol. **27**, 2682–2690 (2010)
9. Burkey, J.: A non-parametric monotonic trend test computing Mann-Kendall Tau, Tau-b, and Sen's Slope written in Mathworks-MATLAB implemented using matrix rotations (2006)
10. Chou, H.-H., Chiu, H.-C., Delaney, N.F., Segrè, D., Marx, C.J.: Diminishing returns epistasis among beneficial mutations decelarates adaptation. Science **322**, 1190–1192 (2011)
11. Costanzo, M.S., Brown, K.M., Hartl, D.L.: Fitness trade-offs in the evolution of dihydrofolate reductase and drug rsistance in *Plasmodium falciparum*. PLoS ONE **6**, e19636 (2011)
12. Crawford, L., Zeng, P., Mukherjee, S., Zhou, X.: Detecting epistasis with the marginal epistasis test in genetic mapping studies of quantitative traits. PLOS Genet. **13**, e1006869 (2017)
13. Crona, K., Gavryushkin, A., Greene, D., Beerenwinkel, N.: Inferring genetic interactions from comparative fitness data. eLife **6**, e28629 (2017)
14. da Silva, J., Coetzer, M., Nedellec, R., Pastore, C., Mosier, D.E.: Fitness epistasis and constraints on adaptation in a human immunodeficiency virus type 1 protein region. Genetics **185**, 293–303 (2010)
15. de Visser, J.A.G.M., Krug, J.: Empirical fitness landscapes and the predictability of evolution. Nat. Rev. Genet. **15**, 480–490 (2014)
16. de Visser, J.A.G.M., Park, S.-C., Krug, J.: Exploring the effect of sex on empirical fitness landscapes. Am. Nat. **174**, S15–S30 (2009)
17. Dean, A.M., Thornton, J.W.: Mechanistic approaches to the study of evolution: the functional synthesis. Nat. Rev. Genet. **8**, 675–688 (2007)
18. Eshel, I., Feldman, M.W.: On the evolutionary effect of recombination. Theor. Popul. Biol. **1**, 88–100 (1970)
19. Fisher, R.A.: The genetical theory of natural selection. Clarendon Press, Oxford (1930)
20. Flynn, K.M., Cooper, T.F., Moore, F.B.G., Cooper, V.S.: The environment affects epistatic interactions to alter the topology of an empirical fitness landscape. PLOS Genet. **9**, e1003426 (2013)
21. Hall, D.W., Agan, M., Pope, S.C.: Fitness epistasis among 6 biosynthtic loci in the budding yeast *Saccharomyces cervisiae*. J. Hered. **1010**, S75–S84 (2010)
22. Heckendorn, R.B., Whitley, D.: Predicting epistasis from mathematical models. Evol. Comput. **7**, 69–101 (1997)
23. Heckendorn, R.B., Wright, A.H.: Efficient linkage discovery by limited probing. Evol. Comput. **12**, 517–545 (2004)
24. Holm, S.: A simple sequentially rejective multiple test procedure. Scand. J. Stat. **6**, 65–70 (1979)
25. Iwasa, Y., Michor, F., Nowak, M.A.: Stochastic tunnels in evolutionary dynamics. Genetics **166**, 1571–1579 (2004)
26. Jiang, P.-P., Corbett-Detig, R.B., Hartl, D.L., Lozovsky, E.R.: Accessible mutational trajectories for the evolution of pyrimethamine resistance in the malaria parasite *Plasmodium vivax*. J. Mol. Evol. **77**, 81–91 (2013)
27. Kauffman, S.A., Weinberger, E.D.: The NK model of rugged fitness landscapes and its application to maturation of the immune response. J. Theor. Biol. **141**, 211–245 (1989)
28. Kendall, M.G.: A new measure of rank correlation. Biometrika **30**, 81–93 (1938)
29. Khan, A.I., Dinh, D.M., Schneider, D., Lenski, R.E., Cooper, T.F.: Negative epistasis between beneficial mutations in an evolving bacterial population. Science **332**, 1193–1196 (2011)

30. Knies, J.L., Cai, F., Weinreich, D.M.: Enzyme efficiency but not thermostability drives cefotaxime resistance evolution in TEM-1 β-lactamase. Mol. Biol. Evol. **34**, 1040–1054 (2017)
31. Kondrashov, A.S.: Deleterious mutations and the evolution of sex. Nature **336**, 435–440 (1988)
32. Leem, S., Jeong, H.-H., Lee, J., Wee, K., Sohn, K.-A.: Fast detection of high-order epistatic interactions in genome-wide association studies using information theoretic measure. Comput. Biol. Chem. **50**, 19–28 (2014)
33. Lozovsky, E.R., Chookajorn, T., Brown, K.M., Imwong, M., Shaw, P.J., Kamchonwongpaisan, S., Neafsey, D.E., Weinreich, D.M., Hartl, D.L.: Stepwise acquisition of pyrimethamine resistance in the malaria parasite. Proc. Natl. Acad. Sci. **106**, 12025–12030 (2009)
34. Lunzer, M., Miller, S.P., Felsheim, R., Dean, A.M.: The biochemical architecture of an ancient adaptive landscape. Science **310**, 499–501 (2005)
35. Malcolm, B.A., Wilson, K.P., Matthews, B.W., Kirsch, J.F., Wilson, A.C.: Ancestral lysozymes reconstructed, neutrality tested, and thermostability linked to hydrocarbon packing. Nature **345**, 86–89 (1990)
36. Maynard Smith, J.: Natural selection and the concept of a protein space. Nature **225**, 563–565 (1970)
37. McCandlish, D.M.: Visualizing fitness landscapes. Evolution **65**, 1544–1558 (2011)
38. Meini, M.-R., Tomatis, P.E., Weinreich, D.M., Vila, A.J.: Quantitative description of a protein fitness landscape based on molecular features. Mol. Biol. Evol. **32**, 1774–1787 (2015)
39. Mira, P.M., Meza, J.C., Nandipati, A., Barlow, M.: Adaptive landscapes of resistance genes change as antibiotic concentrations change. Mol. Biol. Evol. **32**, 2707–2715 (2015)
40. Neher, R.A., Shraiman, B.I.: Competition between recombination and epistasis can cause a transition from allele to genotype selection. Proc. Natl. Acad. Sci. **106**, 6866–6871 (2009)
41. Neidhart, J., Szendro, I.G., Krug, J.: Exact results for amplitude spectra of fitness landscapes. J. Theor. Biol. **332**, 218–227 (2013)
42. Nelson, R.M., Kierczak, M., Carlborg, Ö.: Higher Order Interactions: Detection of Epistasis Using Machine Learning and Evolutionary Computation. In: Gondro, C., van der Werf, J., Hayes, B. (eds.) Genome-Wide Association Studies and Genomic Prediction, pp. 499–518. Humana Press, Totowa (2013)
43. O'Maille, P.E., Malone, A., Dellas, N., Hess Jr., B.A., Smentek, L., Sheehan, I., Greenhagen, B.T., Chappell, J., Manning, G., Noel, J.P.: Quantitative exploration of the catalytic landscape separating divergent plant sesquiterpene synthases. Nat. Chem. Biol. **4**, 617–623 (2008)
44. Orr, H.A.: Fitness and its role in evolutionary genetics. Nat. Rev. Genet. **10**, 531–539 (2009)
45. Palmer, A.C., Toprak, E., Baym, M., Kim, S., Veres, A., Bershtein, S., Kishony, R.: Delayed commitment to evolutionary fate in antibiotic resistance fitness landscapes. Nat. Commun. **6**, 7385 (2015)
46. Phillips, P.C.: The language of gene interaction. Genetics **149**, 1167–1171 (1998)
47. Phillips, P.C.: Epistasis—the essential role of gene interactions in the structure and evolution of genetic systems. Nat. Rev. Genet. **9**, 855–867 (2008)
48. Poelwijk, F., Kiviet, D.J., Tans, S.J.: Evolutionary potential of a duplicated repressor-operator pair: simulating pathways using mutational data. PLoS Comput. Biol. **2**, e58 (2006)
49. Poelwijk, F., Kiviet, D.J., Weinreich, D.M., Tans, S.J.: Empirical fitness landscapes reveal accessible evolutionary paths. Nature **445**, 383–386 (2007)
50. Poelwijk, F.J., Krishna, V., Ranganathan, R.: The context-dependence of mutations: a linkage of formalisms. PLoS Comput. Biol. **12**, e1004771 (2016)
51. Provine, W.B.: Sewall Wright and Evolutionary Biology. University of Chicago Press, Chicago (1986)
52. Rowe, W., Platt, M., Wedge, D.C., Day, P.J., Kell, D.B., Knowles, J.: Analysis of a complete DNA-protein affinity landscape. J. R. Soc. Interface **7**, 397–408 (2010)
53. Sackton, T.B., Hartl, D.L.: Genotypic context and epistasis in individuals and populations. Cell **166**, 279–287 (2016)
54. Sailer, Z.R., Harms, M.J.: Detecting high-order epistasis in nonlinear genotype-phenotype maps. Genetics **205**, 1079–1088 (2017)
55. Sailer, Z.R., Harms, M.J.: High-order epistasis shapes evolutionary trajectories. PLOS Comput. Biol. **13**, e1005541 (2017)
56. Sokal, R.R., Rohlf, F.J.: Biometry. W.H. Freeman and Company, New York (1995)
57. Stadler, P.F.: Landscapes and their correlation functions. J. Math. Chem. **20**, 1–45 (1996)
58. Stadler, P.F.: Spectral Landscape Theory. In: Crutchfield, J.P., Schuster, P. (eds.) Evolutionary Dynamics: Exploring the Interplay of Selection, Accident, Neutrality, and Function, pp. 221–272. Oxford University Press, Oxford (2003)
59. Stadler, P.F., Happel, R.: Random field models for fitness landscapes. J. Math. Biol. **38**, 435–478 (1999)
60. Starr, T.N., Thornton, J.W.: Epistasis in protein evolution. Protein Sci. **25**, 1204–1218 (2016)
61. Sun, X., Lu, Q., Mukherjee, S., Crane, P.K., Elston, R., Ritchie, M.D.: Analysis pipeline for the epistasis search – statistical versus biological filtering. Frontiers in Genetics 5, (2014)

62. Szendro, I.G., Schenk, M., Franke, J., Krug, J., de Visser, J.A.G.M.: Quantitative analyses of empirical fitness landscapes. J. Stat. Mech. **P01**, 005 (2013)
63. Tan, L., Serene, S., Chao, H.X., Gore, J.: Hidden randomness between fitness landscapes limits reverse evolution. Physical Review Letters **106**, 198102 (2011)
64. Weinberger, E.D.: Fourier and Taylor series on fitness landscapes. Biol. Cybern. **65**, 321–330 (1991)
65. Weinreich, D.M., Chao, L.: Rapid evolutionary escape by large populations from local fitness peaks is likely in nature. Evolution **59**, 1175–1182 (2005)
66. Weinreich, D.M., Delaney, N.F., DePristo, M.A., Hartl, D.L.: Darwinian evolution can follow only very few mutational paths to fitter proteins. Science **312**, 111–114 (2006)
67. Weinreich, D.M., Lan, Y., Wylie, C.S., Heckendorn, R.B.: Should evolutionary geneticists worry about high order epistasis? Curr. Opin. Dev. Genet. **23**, 700–707 (2013)
68. Weinreich, D.M., Watson, R.A., Chao, L.: Perspective: sign epistasis and genetic constraint on evolutionary trajectories. Evolution **59**, 1165–1174 (2005)
69. Weirauch, M.T., Yang, A., Albu, M., Cote, A.G., Montenegro-Montero, A., Drewe, P., Najafabadi, H.S., Lambert, S.A., Mann, I., Cook, K., Zheng, H., Goity, A., van Bakel, H., Lozano, J.C., Galli, M., Lewsey, M.G., Huang, E., Mukherjee, T., Chen, X., Reece-Hoyes, J.S., Govindarajan, S., Shaulsky, G., Walhout, A.J.M., Bouget, E.Y., Ratsch, G., Larrondo, L.E., Ecker, J.R., Hughes, T.R.: Determination and inference of eukaryotic transcription factor sequence specificity. Cell **158**, 1431–1443 (2014)
70. Whitlock, M.C., Bourguet, D.: Factors affecting the genetic load in Drosophila: synergistic epistasis and correlations among fitness components. Evolution **54**, 1654–1660 (2000)
71. Wolf, J.B., Brodie, E.D.I., Wade, M.J. (eds.): Epistasis and the Evolutionary Process. Oxford University Press, New York (2000)
72. Wright, S.: Evolution in Mendelian populations. Genetics **16**, 97–159 (1931)
73. Wright, S.: The roles of mutation, inbreeding, crossbreeding and selection in evolution. In: D.F. Jones (eds.) Proceedings of the Sixth International Congress of Genetics, pp. 356–366. Brooklyn Botanic Garden, Menasha (1932)