View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Harvard University - DASH

DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD
DASH.HARVARD.EDU

HARVARD LIBRARY
Office for Scholarly Communication

What can we learn from fitness landscapes?

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters

Citation	Hartl, Daniel L. 2014. "What Can We Learn from Fitness Landscapes?" Current Opinion in Microbiology 21 (October): 51–57. doi:10.1016/j.mib.2014.08.001.
Published Version	10.1016/j.mib.2014.08.001
Citable link	http://nrs.harvard.edu/urn-3:HUL.InstRepos:22898356
Terms of Use	This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Open Access Policy Articles, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#OAP

Elsevier Editorial System(tm) for Current Opinion in Microbiology

Manuscript Draft

Manuscript Number:

Title: What Can We Learn From Fitness Landscapes?

Article Type: 22 Growth&Develop: prokaryotes (2014

Corresponding Author: Dr. Daniel Hartl,

Corresponding Author's Institution:

First Author: Daniel Hartl

Order of Authors: Daniel Hartl

Manuscript
Click here to view linked References

What Can We Learn From Fitness Landscapes?

Daniel L. Hartl

Department of Organismic and Evolutionary Biology

Harvard University

Cambridge, Massachusetts 02138 USA

Contact information: Email: dhartl@oeb.harvard.edu, TEL: 617-396-3917

In evolutionary biology, the fitness landscape of set of mutants is the mapping of genotypes onto phenotypes when the phenotype is fitness or some proxy for fitness such as growth rate or drug resistance. When the set of mutants is not too large, it is possible to create every possible combination of mutants and map these to fitness. Such combinatorially complete datasets have great potential to inform us about molecular and population genetic mechanisms that drive evolutionary change. They indicate how many evolutionary pathways are present in the landscape in which each successive mutational step results in increasing fitness. They also reveal patterns of interaction or epistasis among the mutant sites and whether particular combinations of mutants interact synergistically or antagonistically. Here we examine what has been accomplished already and what it means, but more importantly on what opportunities the approach has opened that have yet to be explored.

The experimental protocol

Given a relatively small number of mutations in the same or different genes that contribute to adaptive evolution, one could construct all possible combinations of the mutations and assay the contribution of each combination of mutants to the adaptation. If there are n genetic changes in the adaptation, with two choices for each, then there are 2^n different combinations. This set of mutations is said to be *combinatorially complete* [1, 2••]. The usual experimental assay for level of adaptation is fitness or some proxy for fitness under specified environmental conditions. Proxies for fitness include growth rate, enzyme activity, and protein stability. In this context, one combination of mutants is regarded as superior to another if the combination increases organismal fitness. Among the $n!$ irreversible pathways (or trajectories) through 2^n combinations of n mutants, a pathway through the sequence space is considered permissible if and only if each step in the pathway increases organismal fitness. Typically, only a limited number of trajectories

through sequence space is permissible [3-8••]. The mapping between genotypes and fitness (or a proxy for fitness) defines the *adaptive topography* for that set of mutants under the given set of conditions. The adaptive topography (or landscape) is a venerable metaphor in evolutionary genetics dating back to Haldane [9] and [10] (see Ref. [11] for review).

One great advantage of combinatorial completeness is that it uncovers the effect of each individual mutation when present in every possible genetic background and hence reveals quantitatively the extent of interaction between pairs, triplets, and higher-order combinations [1,2••]. The approach affords an opportunity to compare actual levels of gene interaction with predicted levels based on systems models of metabolism and reveals tradeoffs between enzyme kinetic parameters, protein stability, and other biochemical and biophysical properties [4,5,12]. Knowing the adaptive topography also enables computer simulations to estimate number and relative probabilities of different evolutionary trajectories [4].

The approach also has limitations. Although it enables estimation of growth rate, metabolic flux, enzyme activity, and other phenotypic characteristics to a high level of accuracy because of replication under controlled, reproducible conditions, the adaptive topography is defined only for that set of conditions and it is not in general known how robust adaptive topographies may be to changing environments. A second limitation is how many genotypes can be constructed and assayed with sufficient replication. For example, Salverda et al. [13] list 18 amino acid residues in TEM-1 β-lactamase at which one or more replacements have a measurable effect on antibiotic resistance in clinical isolates. A combinatorially complete set of these amino acid replacements would require analysis of a prohibitively large number of alleles. The large number of residues that can

contribute to resistance in TEM-1 β-lactamase makes one wonder how an adaptive topography based on a small subset of such mutants might differ according to the TEM-1 sequence background. Nevertheless, some information can be gleaned from combinatorially incomplete data [13,14].

Combinatorially complete datasets

Experiments analyzing combinatorially complete sets of alleles have been summarized by Weinreich et al. [2••] about as well as they can be summarized, and there is no need to repeat their summary here. Suffice it to say that the experimental systems are diverse and include 3–7 genes or protein-coding sites (average 4.6). The systems include:

- 3 examples of metabolic enzymes or pathways — avian lysozyme [15], *Escherichia coli* isopropyl malate dehydrogenase [12,16], and *Solinaceae* sesquiterpine synthetase [17];
- 2 examples of other proteins — mammalian glucocorticoid receptor [18] and HIV glycoprotein [19];
- 3 examples of visible mutants — in *Drosophila melanogaster* [20], *Aspergillis niger* [21,22], and *Saccharomyces cerevisiae* [23];
- 2 examples of adaptations in experimental evolution — in *Metholobacterium extorquens* [24] and *E. coli* [14];
- 5 examples of drug targets — dihydrofolate reductase in *E. coli* [25], β-lactamase in *E. coli* [4,6], *Plasmodium falciparum* dihydrofolate reductase transgenes in *E. coli* [5], *P. falciparum* dihydrofolate reductase transgenes in *S. cervisiae* [7,26,27], and *P. vivax* dihydrofolate reductase transgenes in *S. cerevisiae* and *E. coli* [28•].

From these 15 exemplars as well as other related experiments and observations, one can draw some inferences on the nature of evolution in complex systems. Some of the

inferences are supported by numerous observations and are likely of general applicability, others are supported less well and should be considered tentative.

Inferences so far

1. The number of mutational paths through sequence space is limited and often a relatively small fraction of the theoretical possibilities [4-8••]. This is one of the most strongly supported conclusions among the studies carried out so far.

2. Pathways through sequence space are limited largely by *sign epistasis*, in which a path is inaccessible because one or more steps would entail a decrease in fitness [4,6,8••].

3. Negative pairwise epistasis between beneficial mutations entails a pattern of diminishing returns, in which favorable mutations brought together in combination are less fit than would be expected from their individual effects [24,29].

4. Negative pairwise epistasis for fitness arises because the mapping from biochemical and physiological traits to fitness is nearly always concave [12,14,16,24,30,31]. This pattern is observed experimentally and also one expected on theoretical grounds [29,32••,33]. If fitness is related to a metabolic flux that converges asymptotically to a plateau as a function of increasing enzyme activity, for example, then the fitness–activity curve is concave (Figure 1). The implication concavity is that, near the origin when fitness is low, the relation between activity and fitness is nearly linear, hence mutations that cause small differences in activity are approximately additive with respect to fitness, and there is negligible epistasis. Likewise on the plateau when fitness is high, but at this level even mutations with quite large effects on activity are approximately additive. On the shoulder between these extremes, however, the curvature implies nonadditive effects

of activity on fitness, and mutations in this range are expected to show sign epistasis of the diminishing-returns type.

5. Adaptive reversions are possible in which a favorable substitution incorporated early in a pathway becomes unfavorable and is reversed at a later stage [8,34], but see also Ref. [6]. Adaptive reversions allow indirect routes to attain fitness peaks that may not be directly accessible.

6. Evolutionary pathways often include compensatory mutations that mitigate unfavorable fitness interactions introduced at earlier stages [26,35].

7. While the number of mutational paths through sequence space is constrained, there may nevertheless be enough alternative mutational pathways that the predictability and repeatability of evolutionary trajectories is limited [21].

8. *Reciprocal sign epistasis*, in which single mutants each have a lower fitness than either the double mutant or wildtype, does occur [27,36] but is not pervasive among amino acid replacements [4,5,12,15-19]. The hedge “among amino acid replacements” is important because reciprocal sign epistasis is widespread in RNA molecules that form foldback structures because single mutants that disrupt base pairing in the stem have lower fitness than the double mutant that restores the ability to base pair. In one example of a plant RNA virus, more than half of all significant epistatic interactions were cases of reciprocal sign epistasis [37].

9. Because reciprocal sign epistasis is less prevalent among amino acid replacements than might be expected, fitness landscape can be rugged but are nearly always smoother

than expected were fitness effects of single mutants and their combinations uncorrelated [11,21-23,25]. The fitness effects of alleles that share mutations are correlated for reasons similar to those that explain why offspring resemble their parents.

10. The use of alternating antibiotics that have the same target can restore susceptibility to antibiotics after resistance has evolved. The antibiotics may be structurally similar as in the case of TEM β-lactamase [8••] or structurally dissimilar as in alternate drugs targeting the chloroquine resistance transporter in *P. falciparum* as well alternate drugs targeting dihydroorotate dehydrogenase in *P. falciparum* [38••]. These results are based on laboratory experiments, however clinical data on antibiotic resistance is so far consistent with evolutionary trajectories predicted from in vitro results [5,7,27].

11. Genetic recombination does little to speed adaptation [21-23]. This effect occurs owing to the recombinational breakdown of genotypes on fitness peaks, and it is most pronounced for relatively weak linkage. Theoretically, for tight linkage the situation is reversed [39•]. In principle, in genomes with tight physical linkage between strongly epistatic mutations, recombination allows the attainment of higher fitness peaks owing to the generation of combinations of mutations that may include individually deleterious mutations that enable jumping across fitness valleys [39•].

Next level challenges and opportunities

Despite the impressive list of inferences that have already emerged from combinatorially complete experiments, a number of important issues remain unresolved and some have barely been addressed. While far from exhaustive, the following list highlights some issues that seem to us to follow naturally from the pioneering work already done.

How important is higher-order epistasis? For a combinatorially complete set of n mutant sites or alleles, there are n main effects, $n(n-1)/2$ pairwise epistatic interactions, $n(n-1)(n-2)/3!$ three-way epistatic interactions and, in general, $\binom{n}{k}$ k-way epistatic interactions. Evolutionary geneticists usually limit their considerations to main effects and pairwise interactions, however higher-order interactions ($k \geq 3$) might also be important if for no other reason than because there are so many of them. Weinreich et al. [2••] have estimated levels of higher-order epistasis using Walsh coefficients [40,41], which are linear combinations of fitness values that isolate the effect of each combination of mutants, averaged across all genetic backgrounds, in such a way that each epistatic contribution is independent of all others. For all of the combinatorially complete datasets described above, substantial levels of higher-order epistasis are observed [2••].

Some of the higher-order epistasis is due to the pervasiveness of diminishing-returns epistasis. In the fitness-activity relation in Figure 1, for example, suppose the red dot represents a nonmutant allele, the blue dot any of three single-mutants, the green dot any pairwise combination of the three single mutants, and the orange dot the three-way combination. In this situation, the pairwise and three-way values of epistasis based on Walsh coefficients have the same order of magnitude as the main effects of the alleles. The magnitude of the epistatic effects has to do mainly with the degree of curvature. The effect is smaller in the nearly linear portions of the curve when fitness is ascending or when it has plateaued.

A more traditional way to estimate higher-order epistasis would be through least squares, an approach that automatically tends to maximize the main and second-order

effects and to minimize higher-order effects. When epistasis is estimated by means of least squares for the alleles in Figure 1, for example, the second- and third-order epistatic effects are an order of magnitude smaller than the main effects. A limitation of this approach is that the different orders of epistasis are not independent as they are when using Walsh coefficients.

No matter how higher-order epistasis is estimated, however, the error variance of the estimates is in need of careful investigation. The variance of an estimate of a *k*-way epistatic coefficient may include sums or differences of up to $\binom{n}{k}$ fitness estimates, hence its variance can substantially exceed the average variance of any one fitness estimate.. To the extent that the fitness estimates may be correlated, the variance of the higher-order epistasis may be inflated further.

There is likely no universally best way to measure epistasis, as the best measure of epistasis depends on why it is being measured. For example, one approach may be best for predicting long-term evolutionary outcomes, while another may be best for assessing the forces that drive short-term allele-frequency change in a heterogeneous population. There is even a case to be made for focusing qualitatively on fitness ranks instead of their quantitative values [42-44•]. Some features of fitness landscapes, such as number of local fitness peaks and number of paths to any given peak, lend themselves to this approach. A qualitative approach commends itself because fitness ranks can often be determined more reliably than precise magnitudes. Figure 2 shows a fitness graph with three ordered sites in which red represents mutant sites. The arrows are oriented with the head pointing to the allele associated with the higher fitness. Starting with the all-blue allele, there are two (and only two) accessible paths to the maximum all-red allele, which are indicated by the red arrows. When quantitatively only diminishing returns

epistasis occurs, then the mutants contribute additively to fitness rank and second- and higher-order epistasis disappears. For more complex assignments of fitness rank, second- and higher-order epistasis remains and can be estimated. Analysis of fitness by rank is therefore one way to identify epistatic interactions more complex than those of diminishing returns. The whole question of which is the best measure of second and higher-order epistasis for any specified purpose is rich in possibilities for theoretical analysis.

Are inferences from adaptive landscapes of fitness also valid for other traits? And how do adaptive landscapes of fitness related to lower-level cellular and molecular traits, especially the biochemistry and biophysics of proteins? Fitness is the quintessential higher-order trait, and perhaps landscapes of quantitative traits that are closer to the molecular and cellular level are smoother than those of fitness. This is the case for enzyme thermodynamic stability. Wylie and Shakhnovich [32••] have analyzed a model in which mutations of small effect contribute additively to thermodynamic instability whereas fitness depends on the fraction of molecules present in their folded state. The resulting fitness–stability curve is concave, and mutants affecting protein stability additively show negative epistasis for fitness.

How many other quantitative traits are more nearly additive when measured on an appropriate scale? This is an open question, but it is of critical importance for evaluating risk in complex diseases affected by multiple risk factors. What is the cumulative risk in genomes that include multiple risk factors for hypertension? Type 2 diabetes? Bipolar disorder?

How do adaptive landscapes of fitness change with changing environments? And for antibiotics and other drugs, how do adaptive landscapes compare across related perturbagens? Little data of this sort exist among the combinatorially complete datasets, and even limited data exist outside combinatorial completeness that would allow the key issue to be addressed. One relevant example concerns beneficial mutations in a single-stranded DNA bacteriophage, in which negative pairwise epistasis maintained the same pattern across temperature but intensified as temperature increased [31].

An especially interesting class of environmental agents are small molecules that perturb cellular metabolism (*perturbagens*), such as antibiotics. To what extent do patterns of second- and higher-order epistasis change across a series of chemically related antibiotics that have the same target and act in the same way (e.g., as competitive inhibitors)? Comprehensive data are available only for combinations of β-lactamase mutants when tested against different β-lactam antibiotics [6,8••,13] and combinations of malaria-parasite dihydrofolate reductase mutants against two antimalarial antifolates [7,26,27]. More limited data pertain to paired inhibitors in which one drug is effective against the wildtype allele but not against mutants while the other is effective against mutant alleles but not against wildtype [38••]. These few examples seem to suggest that the adaptive landscape can change quite drastically even for chemically closely related perturbagens, however this conclusion may be misleading because of experimental bias. In all cases studied so far, the chemically related perturbagens were chosen for clinical use or experimental study precisely because the various forms were known to act differently on wildtype and mutant alleles of the drug target. What is needed to assess the robustness of fitness landscapes are studies of chemically related perturbagens that have been chosen with no foreknowledge of their effects on target alleles. Such studies

are likely to be informative from the standpoint of molecular evolution and could be a valuable tool for drug discovery and deployment.

To what extent do orthologous amino replacements in orthologous proteins exhibit similar evolutionary landscapes? Or, to put the question in another way, do orthologous proteins evolve resistance to antibiotics and other drugs through the same or similar amino acid replacements? These questions have hardly been explored [28•] but are of some general interest in revealing whether orthologous proteins can be expected to evolve in parallel pathways when subjected to similar selection pressures. The scarcity of experimental studies of this issue may reflect the fact that it is easy to reduce the questions to absurdity. On the other hand if one studies orthologs that are too different in sequence, then their folding pathways, intrinsically disordered regions, and active-site contacts may be so dissimilar that parallel evolutionary paths could hardly be expected. On the other hand, if one studies orthologs that are virtually identical in sequence, then parallel evolutionary paths are almost assured. The interesting question is at what level of divergence orthologs are still similar enough to be unambiguously aligned but different enough that orthologous mutants might have different biochemical or biophysical properties. The broader question is the extent to which the adaptive landscape is affected by differences in amino acid sequence that do not directly participate in substrate binding or catalytic activity, but may play an essential role in protein folding and the proper orientation of residues in and around the active site. The answer to this question would help in knowing when key resistance residues in one pathogenic species could be used for surveillance to detect emerging resistance in related pathogens, as well as in making best use of evolutionary principles in protein engineering [45••].

Acknowledgements

Many thanks to Daniel M. Weinreich for his characteristically penetrating and helpful comments on the manuscript, and to him and Christopher L. Hartl for their clarifying comments and explanations of the use of Walsh coefficients to quantify epistasis. Special thanks to Elena R. Lozovsky for pointing out ambiguities and other trouble spots. This research was supported by U.S. National Institutes of Health grant AI106734.

Regrettably, many interesting papers could not be included owing to limitations of space.

References and recommended reading

Papers of particular interest published within the period of review have been highlighted as:

• of special interest

•• of outstanding interest

1. Poelwijk FJ, Kiviet DJ, Weinreich DM, Tans SJ: **Empirical fitness landscapes reveal accessible evolutionary paths**. *Nature* 2007, **445**:383-386.

••2. Weinreich DM, Lan Y, Wylie CS, Heckendorn RB: **Should evolutionary geneticists worry about higher-order epistasis?** *Curr Opin Genet Dev* 2013, **23**:700-707.

A "must-read" analysis proposing a measure of higher-order epistasis based on Walsh coefficients that is quite distinct from the classical Fisherian least-squares approach.

3. Weinreich DM, Watson RA, Chao L: **Perspective: Sign epistasis and genetic constraint on evolutionary trajectories**. *Evolution* 2005, **59**:1165–1174.

4. Weinreich DM, Delaney NF, DePristo MA, Hartl DL: **Darwinian evolution can follow only very few mutational paths to fitter proteins**. *Science* 2006, **312**:111-114.

5. Lozovsky ER, Chookajorn T, Brown KM, Imwong M, Shaw PJ, Kamchonwongpaisan S, Neafsey DE, Weinreich DM, Hartl DL: **Stepwise acquisition of pyrimethamine resistance in the malaria parasite**. *Proc. Natl. Acad. Sci. U. S. A.* 2009, **106**:12025-12030.

6. Tan L, Serene S, Chao HX, Gore J: **Hidden randomness between fitness landscapes limits reverse evolution**. *Phys. Rev. Lett.* 2011, **106**:198102.

7. Costanzo MS, Hartl DL: **The evolutionary landscape of antifolate resistance in Plasmodium falciparum**. *J. Genet.* 2011, **90**:187-190.

••8. Goulart CP, Mahmudi M, Crona KA, Jacobs SD, Kallmann M, Hall BG, Greene DC, Barlow M: **Designing antibiotic cycling strategies by determining and understanding local adaptive landscapes**. *PLoS One* 2013, **8**:e56040.

A study of the fitness landscapes of TEM β-lactamase alleles with respect to 15 β-lactam antibiotics showing that some compounds are able to inhibit alleles that are resistant to structurally similar compounds.

9. Haldane JBS: **A mathematical theory of natural selection, part VIII: Metastable populations**. *Proc. Cambridge Phil. Soc.* 1931, **27**:137-142.

10. Wright S: **The roles of mutation, inbreeding, crossbreeding and selection in evolution**. *Proc. Sixth Int. Cong. Genet.* 1932, **1**:356-366.

11. Carneiro M, Hartl DL: **Adaptive landscapes and protein evolution**. *Proc. Natl. Acad. Sci. USA* 2009, **107 Suppl. 1**:1747-1751.

12. Lunzer M, Miller SP, Felsheim R, Dean AM: **The biochemical architecture of an ancient adaptive landscape**. *Science* 2005, **310**: 499-501.

13. Salverda MLM, de Visser JAGM, Barlow M: **Natural evolution of TEM-1 β-lactamase: experimental reconstruction and clinical relevance**. *FEMS Microbiol. Rev.* 2010, **34**:1015-1036.

14. Khan AI, Dinh DM, Schneider D, Lenski RE, Cooper TF: **Negative epistasis between beneficial mutations in an evolving population**. *Science* 2011, **332**:1193-1196.

15. Malcolm BA, Wilson KP, Matthews BW, Kirsch JF, Wilson AC: **Ancestral lysozymes reconstructed, neutrality tested, and thermostability linked to hydrocarbon packing**. *Nature* 1990, **345**:86-89.

16. Miller SP, Lunzer M, Dean AM: **Direct demonstration of an adaptive constraint**. *Science* 2006, **314**:458-461.

17. O'Maille PE, Malone A, Dellas N, Hess BAJ, Smentek L, Sheehan I, Greenhagen BT, Chappell J, Manning G, Noel JP: **Quantitative exploration of the catalytic landscape separating divergent plant sesquiterpene synthases**. *Nat. Chem. Biol.* 2008, **4**:617-623.

18. Bridgham JT, Carroll SM, Thornton JW: **Evolution of hormoner eceptor complexity by molecular exploitation**. *Science* 2007, **312**:97-100.

19. da Silva J, Coetzer M, Nedellec R, Pastore C, Mosier DE: **Fitness epistasis and constraints on adaptation in a human immunodeficiency virus type 1 protein region**. *Genetics* 2010, **185**:293-303.

20. Whitlock MC, Bourguet D: **Factors affecting the genetic load in Drosophila: synergistic epistasis and correlations among fitness components**. *Evolution* 2000, **54**:1654-1660.

21. Franke J, Klözer A, de Visser JAGM, Krug J: **Evolutionary accessibility of mutational pathways**. *PLOS Comput. Biol.* 2011, **7**:e1002134.

22. de Visser JAGM, Park S-C, Krug J: **Exploring the effect of sex on empirical fitness landscapes**. *Am. Nat.* 2009, **174**:S15-S30.

23. Hall DW, Agan M, Pope SC: **Fitness epistasis among 6 biosynthtic loci in the budding yeast *Saccharomyces cervisiae***. *J. Hered.* 2010, **1010**:S75-S84.

24. Chou H-H, Chiu H-C, Delaney NF, Segrè D, Marx CJ: **Diminishing returns epistasis among beneficial mutations decelerates adaptation**. *Science* 2011, **332**:1190-1192.

25. Aita T, Iwakura M, Hasumi Y: **A cross-section of the fitness landscape of dihydrofolate reductase**. *Protein Eng*. 2001, **14**:633-638.

26. Brown KM, Costanzo M, Xu W, Roy SW, Lozovsky ER, Hartl DL: **Compensatory mutations restore fitness during the evolution of dihydrofolate reductase**. *Mol. Biol. Evol.* 2010, **27**:2682-2690.

27. Costanzo MS, Brown KM, Hartl DL: **Fitness trade-offs in the evolution of dihydrofolate reductase and drug resistance in *Plasmodium falciparum***. *PLoS One* 2011, **6**:e19636.

•28. Jiang PP, Corbett-Detig RB, Hartl DL, Lozovsky ER: **Accessible mutational trajectories for the evolution of pyrimethamine resistance in the malaria parasite *Plasmodium vivax***. *J. Mol. Evol.* 2013, **77**:81-91.

Combinatorially complete analysis of four mutants in an important antimalarial drug target showing how fitness landscapes can change as a function of drug dosage.

29. Schenk MF, Szendro IG, Salverda ML, Krug J, de Visser JA: **Patterns of Epistasis between beneficial mutations in an antibiotic resistance gene**. *Mol. Bio. Evol.* 2013, **30**:1779-1787.

30. Hartl DL, Dykhuizen DE, Dean AM: **Limits of adaptation: The evolution of selective neutrality**. *Genetics* 1985, **111**:655-674.

31. Caudle SB, Miller CR, Rokyta DR: **Environment determines epistatic patterns for a ssDNA virus**. *Genetics* 2014, **196**:267-279.

••32. Wylie CS, Shakhnovich EI: **A biophysical protein folding model accounts for mutational fitness effectgs in viruses**. *Proc. Natl. Acad. Sci. U. S. A.* 2011, **108**:9916-9921.

Combines principles of population genetics with those of protein folding and shows that protein thermodynamic stability can account for a substantial fraction of observed mutational effects in viruses.

33. Chiu HC, Marx CJ, Segrè D: **Epistasis from functional dependence of fitness on underlying traits**. *Proc. Biol. Sci.* 2012, **279**:4156-4164.

34. DePristo MA, Hartl DL, Weinreich DM: **Mutational reversions during adaptive protein evolution**. *Mol. Biol. Evol.* 2007, **24**:1608-1610.

35. DePristo MA, Weinreich DM, Hartl DL: **Missense meandering through sequence space**. *Nature Rev. Genet.* 2005, **6**:678-687.

36. Kvitek DJ, Sherlock G: **Reciprocal sign epistasis between frequently experimentally evolved adaptive mutations causes arugged fitness landscape**. *PLoS Genet*. 2011, **7**:e1002056.

37. Lalić J, Elena SF: **Magnitude and sign epistasis among deleterious mutations in a positive-sense plant RNA virus**. *Heredity (Edinb)*. 2012, **109**:71-77.

••38. Lukens AK, Ross LS, Heidebrecht R, Javier Gamo F, Lafuente-Monasterio MJ, Booker ML, Hartl DL, Wiegand RC, Wirth DF: **Harnessing evolutionary fitness in *Plasmodium falciparum* for drug discovery and suppressing resistance**. *Proc. Nat. Acad. Sci. USA* 2014, **111**:799-804.

A demonstration that screeing large libraries of small molecules can sometimes reveal "paired inhibitors" in which one inhibits wildtype sensitive alleles and the other mutant resistance alleles, with potentially important implications for drug development and deployment.

•39. Watson RA, Weinreich DM, Wakeley J: **Genome structure and the benefit of sex**. *Evolution* 2011, **65**:523-536.

Theoretical analysis exploring the evolution of sexual and asexual populations on multi-peaked fitness landscapes that outlines the conditions in which sexuals can evolve high-fitness genotypes that are unattainable in asexuals

40. Heckendorn RB, Whitley D: **Predicting epistasis from mathematical models**. *Evol. Comput.* 1997, **7**:69-101.

41. Iglesias M, Verschoren A, Vidal TC: **Higher order functions and Walsh coefficients revisited**. *Bull. Belg. Math. Soc. Simon Stevin* 2008, **15**:385-568.

42. Beerenwinkel N, Pachter L, Sturmfels B: **Epistasis and shapes of fitness landscapes**. *Statistica Sinica* 2007, **17**:1317–1342.

43. Beerenwinkel N, Pachter L, Sturmfels B, Elena SF, Lenski RE: **Analysis of epistatic interactions and fitness landscapes using a new geometric approach**. *BMC Evol. Bio.* 2007, **7**:60.

•44. Crona K, Green D, Barlow M: **The peaks and geometry of fitness landscapes**. *J. Theor. Biol.* 2013, **317**:1-10.

Presents a qualitative analysis of fitness landscapes by means of fitness graphs in which nodes represent genotypes and arrows point toward more fit genotypes.

••45. Reynolds KA, Russ WP, Socolich M, Ranganathan R: **Evolution-based design of proteins**. *Methods Enzymol.* 2013, **523**:213-235.

Reviews statistical analyses revealing that interactions among amino acids in naturally occurring proteins are sparse and nested into hierarchies, and suggests how this principle can be taken into account in protein engineering.

Figure legends:

Fig. 1. Mapping of the activity of a hypothetical enzyme onto fitness. The model assumes that fitness is proportion to flux through a metabolic pathway when enzyme activity is limiting to flux according to simple Michaelis-Menten enzyme kinetics, and the curve is normalized to a fitness of 1 when activity equals 25. The colored dots correspond to fitness for wildtype (red), any of three single mutants (blue), any pair of double mutants (green), and thr triple mutant (orange).

Fig. 2. Qualitative analysis of a fitness landscape with three mutant sites or genes (circles). Wildtype is symbolized by blue, mutant by red. Red arrows indicate pathways accessible from the nonmutant genotype (all blue) and blue arrows represent pathways that are inaccessible from this state. The fitness maximum is realized by the triple mutant, and there are two accessible pathways to this state from the double mutant.

Figure 1

Click here to download high resolution image

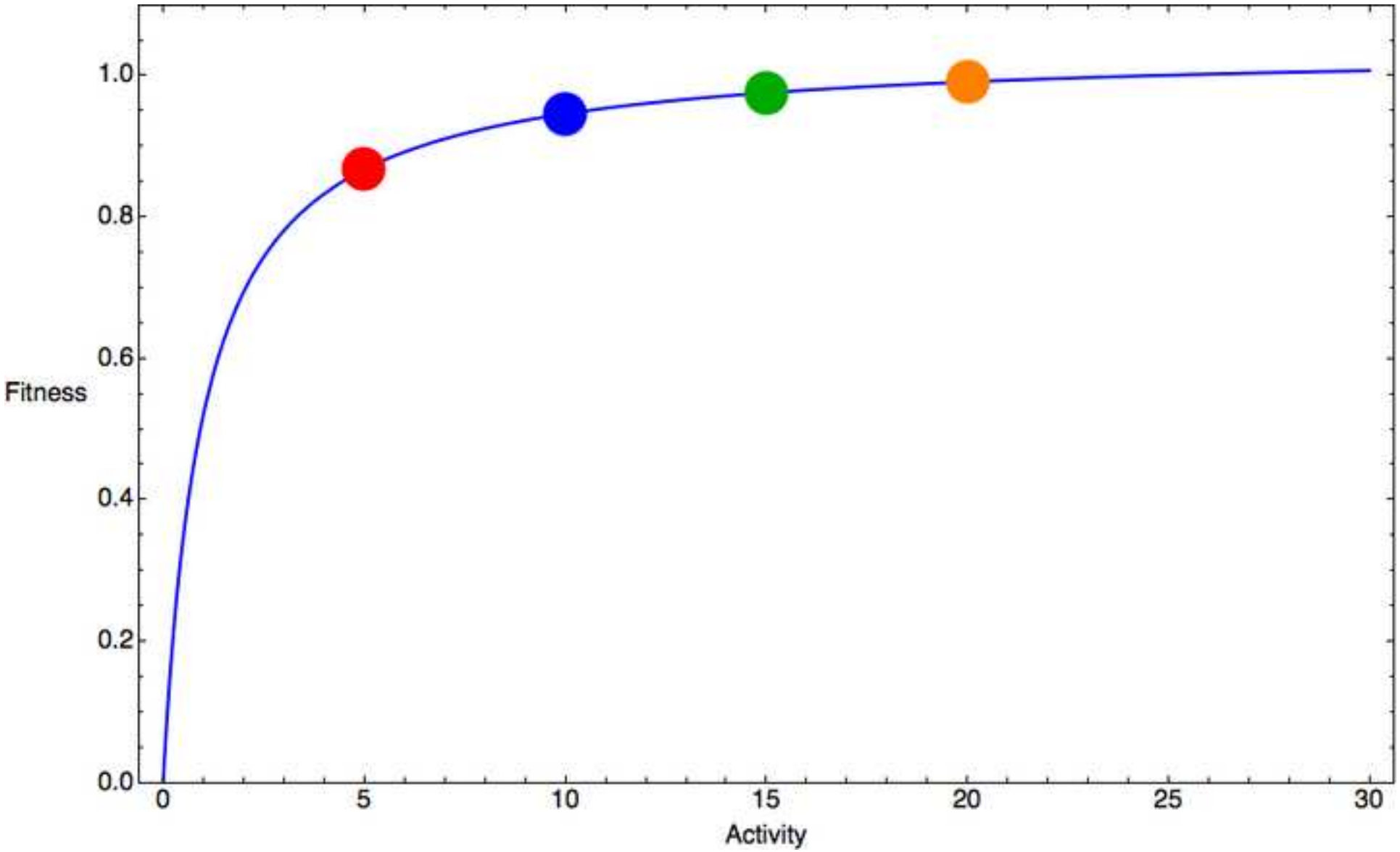
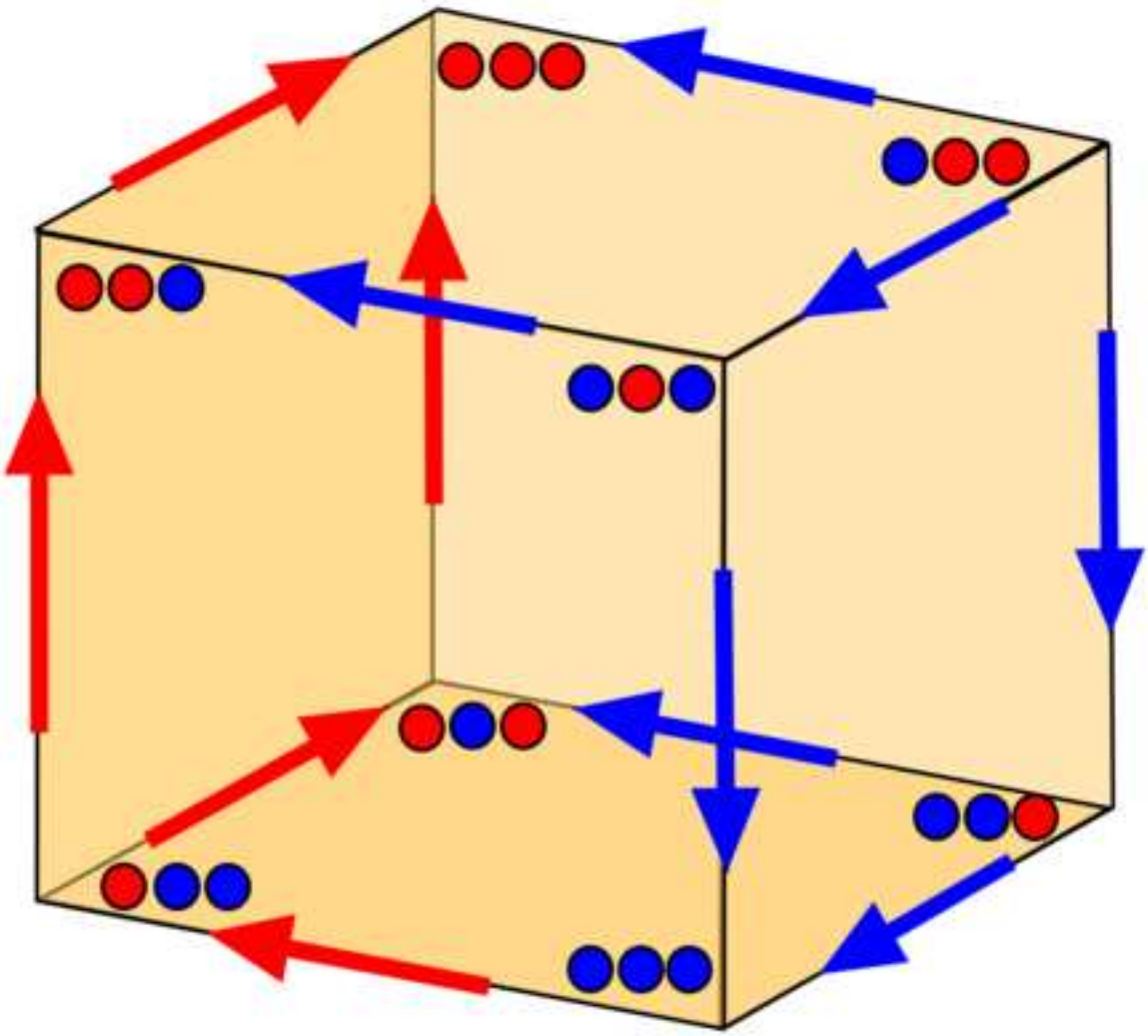


Figure 2

Click here to download high resolution image



***Highlights (for review)**

- Fitness landscapes can help in understanding constraints on evolutionary change
- Combinatorially complete reveal patterns of higher-order fitness interactions
- Actual fitness landscapes are much smoother than random fitness landscapes
- Reciprocal sign epistasis can occur but is not pervasive
- Fitness landscapes open great opportunities for further research