# Adaptive mutation of biochemical reaction constants:
# Fisher's geometrical model without pleiotropy

Ryan N. Gutenkunst[a,*], James P. Sethna[b]

[a]*Theoretical Biology and Biophysics & Center for Nonlinear Studies, Los Alamos National Laboratory, Los Alamos, NM 87545*
[b]*Laboratory of Atomic and Solid State Physics, Cornell University, Ithaca, NY 14853*

## Abstract

The distribution of fitness effects of adaptive mutations remains poorly understood, both empirically and theoretically. We study this distribution using a version of Fisher's geometrical model without pleiotropy, such that each mutation affects only a single trait. We are motivated by the notion of an organism's chemotype, the set of biochemical reaction constants that govern its molecular constituents. From physical considerations, we expect the chemotype to be of high dimension and to exhibit very little pleiotropy. Our model generically predicts striking cusps in the distribution of the fitness effects of arising and fixed mutations. It further predicts that a single element of the chemotype should comprise all mutations at the high-fitness ends of these distributions. Using extreme value theory, we show that the two cusps with the highest fitnesses are typically well-separated, even when the chemotype possesses thousands of elements; this suggests a means to observe these cusps experimentally. More broadly, our work demonstrates that new insights into evolution can arise from the chemotype perspective, a perspective between the genotype and the phenotype.

*Keywords:* fitness landscape; pleiotropy; fitness effects distribution; extreme value theory

## 1. Introduction

Adaptive mutation is fundamental to the evolutionary process, and it is medically important to the emergence of drug resistance in microbes [1] and tumors [2]. Given the selective advantage of a mutation, the probability that it fixes in a population (i.e., rises to frequency 1) and the mean time to do so are well-known [3]. Comparatively little is known, however, about the distribution of selective advantages among new mutations. This distribution can be experimentally measured by confronting genetically identical populations with a novel environment such a new food source and measuring the fitness of newly arising mutations [4]. Such measurements are difficult, because adaptive mutations are rare; thus theoretical analysis can offer important insights [5].

A popular predictive framework for studying adaptive evolution is R. A. Fisher's geometrical model, which considers adaptation in phenotypic "trait" space [6]. Mutations are characterized by the phenotypic changes they induce, which correspond to moves in trait space. Fisher used this model to argue that evolution is primarily driven by the accumulation of many mutations that each have only a small effect [6]. This argument was influential until Motoo Kimura pointed out that mutations with larger effects are more likely to fix, so most adaptive mutations that fix have intermediate effect [7].

Recent studies have applied Fisher's model to a gamut of questions in evolutionary biology and population genetics; these include the distribution of mutation fitness effects near an optimum [8], sequential adaptation [9, 10], and the load of deleterious mutations carried by finite populations [11, 12]. Of particular note, predictions from the model regarding epistasis compare favorably with data [13]. The model predicts a roughly exponential distribution of fitness effects for new mutations [14], similar to mutational landscape models of adaptive evolution [15]. This prediction is consistent with experiments in viruses [16] and bacteria [17], although more recent experiments by Rokyta et al. point toward a truncated distribution [18]. Here we consider a geometrical model without pleiotropy, a model in which each mutation affects only a single trait. We are motivated by considering the phenotype at a finer scale than is typical.

One can view the information specifying an organism through a variety of scales [19]. On the largest scale, the phenotype of the entire organism, a single mutation often affects multiple traits, implying substantial pleiotropy. On the finest scale, the genotype, a single mutation often affects only one amino acid codon or one regulatory binding site, implying no pleiotropy. Systems biology is often modeled at the intermediate scale of biochemical reaction constants; multiple codons combine to determine a single biochemical reaction constant and multiple constants combine to determine a single phenotypic trait. Motivated by this useful intermediate level of description, we introduced the word "chemotype" [19] to refer to the set of biochem-

---

*Corresponding author

ical reaction constants determining the rates of molecular reactions in an organism. Other authors have considered specific biochemical reaction constants to be aspects of the phenotype, for example Hartl, Dykhuizen and Dean [20]. We find it useful to distinguish the chemotype, because it differs in important ways from the large-scale phenotype typically considered.

The chemotype differs from the large-scale phenotype in both dimensionality and pleiotropy. The number of independent high-level phenotypic traits for even a complex organism may be modest [21, 22]. The number of independent elements of a chemotype, however, is comparable to the number of an organism's genes. Each gene codes for a protein or RNA with its own biochemical reaction constants, so each gene contributes at least one element to the chemotype. The chemotype is additionally distinguished by very low pleiotropy, the degree to which single mutations affect multiple traits. Recent experiments on mouse skeletal traits have demonstrated that this system possess a moderate degree of pleiotropy; a given mutation typically affects around five traits [23]. By contrast, a single mutation is expected to affect only one or a few elements of the chemotype. This is because single-nucleotide mutations are dominant in short-term and laboratory evolution [24, 25], and they typically change only a single protein residue or a single DNA binding site. Such a change will in turn impact only one or a few biochemical reaction constants, implying very low pleiotropy in chemotype space.

Other authors have considered zero pleiotropy geometric models in the study of drift load [26, 12]. We focus here on the distributions of fitness effects of adaptive mutations that arise and that subsequently fix in a population. A general argument shows that these distributions possess sharp cusps, one for each element of the chemotype. Given the high dimensionality of chemotype space, however, it is not obvious whether these cusps are observable. To address this question, we study a more specific model, in which the fitness landscape is Gaussian. For this model, we show using extreme value theory that the two cusps with the highest fitnesses are well-separated, even in a space with thousands of dimensions. This suggests that the cusps, and thus the nature of evolution in chemotype space, can be studied experimentally.

## 2. Model

As illustrated in Fig. 1, the state of an organism with $N$ chemotype elements can be represented as a point in $N$-dimensional space: $\vec{k} = \langle k_1, k_2, \ldots k_N \rangle$. The change in state caused by a mutation can be described by an $N$-dimensional vector $\vec{r}$; the mutant has chemotype $\vec{k} + \vec{r}$. Because mutations will typically change only one or a few elements of the chemotype, most pairs of mutations are *orthogonal* in this space. We thus restrict our attention to mutations with zero pleiotropy, which change only a single chemotype element at a time. Thus $\vec{r} = r\,\hat{r}_i$, where $r$ the
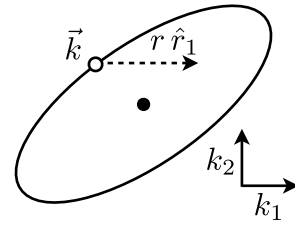


Figure 1: Illustration of the model. Motived by the evolution of biochemical reaction constants, we consider evolution in a high-dimensional space with no pleiotropy (chemotype space). The current uniform chemotype $\vec{k}$ of a population is a point in this space. The optimal chemotype is the origin of our coordinate system and lies at the center of the fitness contours. In the absence of pleiotropy, mutations change one element of $\vec{k}$ at a time, so moves are made along the coordinate axes. The dashed arrow indicates an adaptive mutation of magnitude $r$ in element $k_1$.

size of the mutation (which may be negative) and $\hat{r}_i$ is a unit vector along the $i$th coordinate axis. We define each $\hat{r}_i$ so that mutations which increase fitness have positive values of $r$. We work in the limit of strong selection and weak mutation, so that the population is genetically homogenous aside from rare mutants that arise one at a time and either fix or are lost before the next mutation arises. In this limit, the state of the entire population corresponds to a single point $\vec{k}$ in chemotype space, and fixation of the mutation $\vec{r}$ moves the entire population to $\vec{k} + \vec{r}$.

### 2.1. Gaussian landscape

For analysis, we specialize to a Gaussian fitness landscape in which the fitness $W(\vec{k})$ of a population with chemotype $\vec{k}$ is

$$W(\vec{k}) = \exp\left(-\frac{1}{2}\,\vec{k}\cdot\mathbf{S}\cdot\vec{k}\right), \tag{1}$$

where $\mathbf{S}$ is a symmetric positive definite matrix and $\mathbf{S}\cdot\vec{k}$ denotes the dot product of the matrix $\mathbf{S}$ and the vector $\vec{k}$. Without loss of generality, the optimum fitness is set to one.

It is convenient to work with the logarithmic fitness change $Q$, introduced by Waxman and Welch [27] and defined as

$$Q \equiv \log\left[W(\vec{k}+\vec{r})\,/\,W(\vec{k})\right]. \tag{2}$$

$Q$ is related to the selection coefficient $s$ by $s = e^Q - 1$, and for mutations with small selective advantage $Q \approx s$. Adaptive mutations are those with $Q > 0$. For the Gaussian landscape, the log-fitness change caused by a mutation of size $r$ in chemotype element $i$ is

$$Q_i(r) = -\vec{k}\cdot\mathbf{S}\cdot\hat{r}_i\, r - \frac{1}{2}\,\hat{r}_i\cdot\mathbf{S}\cdot\hat{r}_i\, r^2. \tag{3}$$

The largest possible gain in log-fitness achievable by mutating chemotype element $i$ is denoted $\theta_i$ and obtained by

maximizing $Q_i(r)$ with respect to $r$:

$$\theta_i = \frac{\left|\vec{k}\cdot\mathbf{S}\cdot\hat{r}_i\right|^2}{2\,\hat{r}_i\cdot\mathbf{S}\cdot\hat{r}_i}. \tag{4}$$

The magnitude of the largest possible mutation in chemotype element $i$ that can be made without decreasing fitness is $\rho_i$:

$$\rho_i = 2\frac{\left|\vec{k}\cdot\mathbf{S}\cdot\hat{r}_i\right|}{\hat{r}_i\cdot\mathbf{S}\cdot\hat{r}_i}. \tag{5}$$

These quantities are illustrated in Fig. 2A.

Many of our results are derived for spherically symmetric fitness landscapes, for which $\mathbf{S} = \lambda\mathbf{I}$, where $\mathbf{I}$ is the identity matrix. For such a landscape,

$$\theta_i = \frac{\lambda|\vec{k}\cdot\hat{r}_i|^2}{2} \equiv Q_{tot}|\hat{k}\cdot\hat{r}_i|^2, \tag{6}$$

and

$$\rho_i = 2|\vec{k}\cdot\hat{r}_i|. \tag{7}$$

Here $Q_{tot} \equiv \lambda|\vec{k}|/2 = -\log W(\vec{k})$ is the difference in log-fitness between the optimum (which has a fitness of one) and the current chemotype.

## 3. Results

We first show generally that adaptation on a smooth fitness landscape without pleiotropy leads to cusps in the distribution of fitness effects of newly arising and fixed mutations, one cusp for each element of the chemotype (i.e., each dimension of the landscape). To assess whether these cusps will be observable, we specialize to Gaussian landscapes for which we can perform detailed calculations. We show that the cusps will be difficult to directly observe in a histogram of experimental fitness measurements. Nevertheless, we show that the cusps do have observable consequences, because the two cusps with the highest fitnesses are typically well-separated, even if the landscape has thousands of dimensions and deviates strongly from a sphere.

### 3.1. Cusps

Fig. 2A illustrates the slice $Q_i$ of the fitness landscape accessible by mutations to a particular chemotype element $i$. For a mutation of size $r$, the range of mutations $\Delta r$ about $r$ that produce fitness changes in a given range $\Delta Q$ is inversely proportional to the slope $dQ_i/dr$. By definition, $Q_i$ is at a maximum for the fittest mutation $r_i^*$ and thus has zero slope there. This yields an infinite inverse and thus a cusp in the distribution of fitness effects of adaptive mutations to chemotype element $i$, as illustrated in Fig. 2B.

The above argument relies only on a lack of pleiotropy, a smooth fitness landscape, and a distribution of fitness affects which is non-singular and broad enough to access the optimal mutation of a given chemotype element. A natural question is whether these cusps will be observable in experiments. To address this, we analyze a more specific landscape for which we can perform concrete calculations.

### 3.2. Gaussian landscape

Comparisons between empirical mutation effect distributions in different environments for several organisms support a Gaussian form for the fitness landscape close to the optimum chemotype [28]. For the remainder of this paper, we assume a Gaussian landscape as in Eq. 1.

We must also specify the distribution of mutational effects $f(r)$ on the chemotype. In Appendix A, we use the fact that most mutations are deleterious [29, 30] to show that this distribution must typically span the full range of adaptive mutations for all elements of the chemotype. In other words, mutations must often 'hop over' the ridge of increased fitness. For computational ease, we take $f(r)$ to be uniform over the range of adaptive mutations; our qualitative results are robust to this choice.

Given these two assumptions, we can calculate the distribution $f_a(Q)$ of fitness effects for adaptive mutations, as detailed in Appendix B:

$$f_a(Q) \propto \sum_i \frac{1}{\sqrt{\hat{r}_i\cdot\mathbf{S}\cdot\hat{r}_i}\sqrt{\theta_i - Q}}. \tag{8}$$

Here the sum is over elements of the chemotype, and each element contributes its own cusp. Fig. 3A plots $f_a(Q)$ for a population at a particular random chemotype $\vec{k}$ in a 30-dimensional spherical fitness landscape (in which $\mathbf{S}$ is proportional to the identity matrix). Note that the distribution is bound by a roughly exponential envelope. As detailed in Appendix C, if we average our distribution over initial states $\vec{k}$ with a given fitness, we obtain Waxman and Welch's previous result [27] for the spherical model with maximum pleiotropy.

For a large population, the probability that an adaptive mutation fixes is proportional to its fitness effect $Q$ [7, 31]; thus the density of fitness effects of fixed mutations $f_f(Q)$ is

$$f_f(Q) \propto Q f_a(Q). \tag{9}$$

This density of fixed mutations is shown in Fig. 3B for the same spherical fitness landscape and initial chemotype $\vec{k}$ as in Fig. 3A. The cusps at large $Q$ are much more prominent in the distribution of fixed mutation fitness effects.

Distributions of fitness effects can be measured experimentally by introducing identical populations to identical novel environments and tracking mutations that sweep through them. The histogram in Fig. 3B simulates such an experiment, representing 1000 samples from $f_f(Q)$, each polluted by Gaussian measurement noise in $Q/Q_{tot}$ with a standard deviation 0.01. Given the rarity of adaptive mutations, such an experiment would be difficult, but even
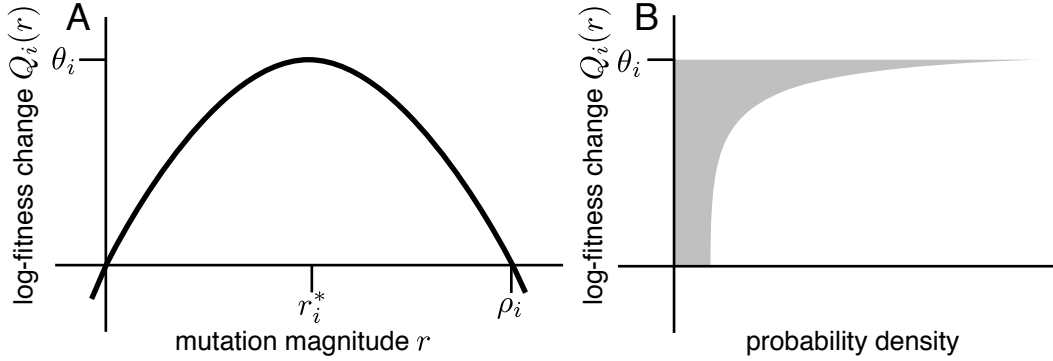
3

Figure 2: Origin of cusps. A: The log-fitness $Q_i(r)$ as a function of the mutation size $r$ in chemotype element $i$. $Q_i(r)$ has a maximum of $\theta_i$ at $r = r_i^*$, and it returns to zero at $r = \rho_i$. B: The probability density of mutation fitness effects. It generically has a cusp at $Q_i(r) = \theta_i$, corresponding to the point at which $Q_i(r)$ has zero slope. The density is plotted sideways to emphasize the connection between the cusp and the maximum of $Q_i(r)$.
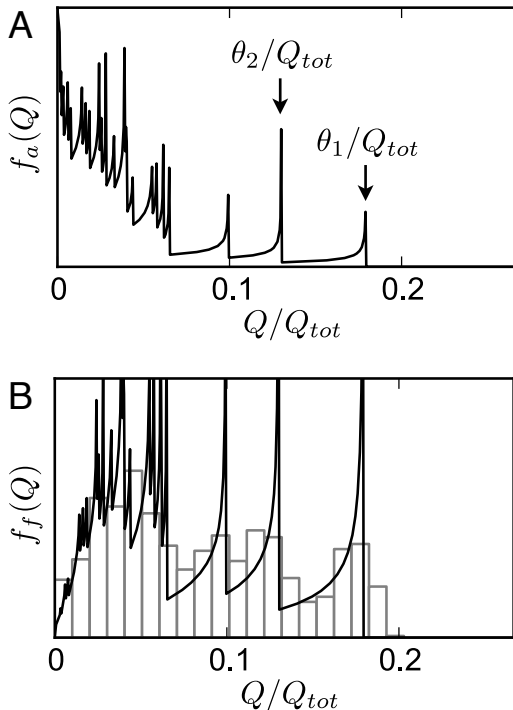


Figure 3: Example of cusps. A: The probability distribution $f_a(Q)$ of the fitness effects of adaptive mutations for a 30-dimensional spherical fitness landscape with a random initial chemotype $\vec{k}$. B: The probability distribution $f_f(Q)$ of fitness effects of fixed mutations, for the same $\vec{k}$ as in A. The cusps at large values of the fitness $Q$ are much more prominent. The grey histogram shows 1000 samples from the distribution, each including a 1% error in the measurement of $Q/Q_{tot}$.

such a difficult experiment cannot directly resolve the cusps. We now show, however, that the spacing between the fittest and second-fittest cusps is typically substantial. This implies that the upper end of the fitness distribution should be dominated by mutations to a single element of the chemotype, a prediction that can be tested experimentally.

### 3.2.1. Cusp spacings

Each cusp in Fig. 3 corresponds to mutations affecting a different chemotype element $k_i$. Thus our model not only predicts cusps, but also predicts that the most adaptive mutations will all affect the *same* element of the chemotype. To experimentally test this prediction, it suffices to measure relative fitness differences of order $\Delta$, where

$$\Delta \equiv (\theta_1 - \theta_2)/\theta_1 \qquad (10)$$

is the normalized separation between the two cusps with the highest fitnesses. In Appendix D we derive the distribution of $\Delta$ predicted by our model for a spherical landscape, using methods of extreme value theory [32].

The solid line in Fig. 4 is the exact asymptotic result (using Eq. D.4 and D.5) for the mean of $\Delta$, given a spherical fitness landscape. The dashed line is the approximation

$$\langle \Delta \rangle \approx \frac{1}{1 + \log N + \log \sqrt{2/\pi}}, \qquad (11)$$

which is valid for large landscape dimension $N$. The circles in Fig. 4 are the results from numerical simulations in the spherical landscape at each $N$. The agreement between the exact asymptotic result and the numerical simulations is excellent, and the approximate result captures the trend well. Note that $\langle \Delta \rangle$ declines very slowly as a function of $N$; for a chemotype with $N = 10,000$ elements the mean $\Delta$ is approximately 0.11, a relative fitness difference that is straightforward to measure experimentally. For comparison, Fig. 3 has $\Delta \approx 0.27$, which is approximately the predicted $\langle \Delta \rangle$ when the chemotype has $N = 30$ elements. Thus our model predicts that, even for a high-dimensional
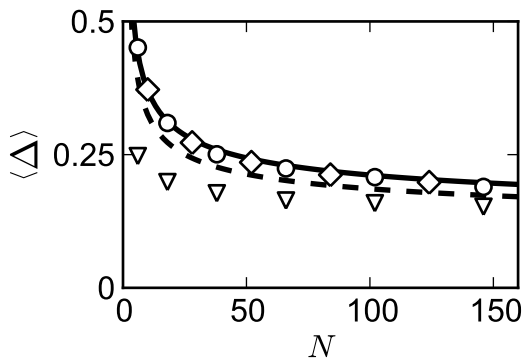
4

Figure 4: Mean relative spacing $\Delta$ between the fittest and second-fittest cusps as a function of landscape dimension $N$. The solid line is the asymptotically exact result for the spherical fitness landscape, while the dashed line is the approximation of Eq. 11. The circles are numerical simulations for the spherical landscape, while the diamonds and triangles are simulation results for mildly and severely non-spherical landscapes, respectively. The mean value of $\Delta$ declines very slowly with $N$ (especially for large $N$), suggesting that the cusps will typically be well-separated even for chemotype spaces of high dimension.

chemotype, a substantial range $\Delta$ of the most adaptive mutations will affect the same element of the chemotype.

### 3.2.2. Non-spherical landscapes

To assess the robustness of our results regarding cusp spacing, we numerically test them in non-spherical landscapes. Studying non-spherical fitness landscapes also implicitly considers different mutation scales amongst chemotype elements, because any differences in the typical size of chemotype mutation effects on different elements (anisotropic $f(r)$) can be eliminated by rescaling the chemotype elements $k_i$. A convenient way to characterize a landscape is by the eigenvalues of $\mathbf{S}$. Spherical landscapes have all eigenvalues equal, while for non-spherical landscapes the width of the fitness contour along any given eigenvector of $\mathbf{S}$ is proportional to the square root of the corresponding eigenvalue, so landscapes with a larger range of eigenvalues are more non-spherical.

For a given landscape $\mathbf{S}$ and initial chemotype $\vec{k}$, $\Delta$ can be calculated numerically from the definition of $\theta_i$. In the tests described below, the mean of $\Delta$ is calculated from $10^4$ simulations, each instance involving an independent landscape $\mathbf{S}$ and initial chemotype $\vec{k}$. The eigenvectors of $\mathbf{S}$ were random orthogonal vectors, and the initial chemotypes were chosen at random among those with a fixed fitness $Q_{tot}$, as described in Appendix E.

The diamonds in Fig. 4 result from mildly non-spherical fitness landscapes corresponding to eigenvalues of $\mathbf{S}$ drawn uniformly from the range $0.4 < \lambda < 3.6$, following Waxman [33]. The deviations of $\langle \Delta \rangle$ from the spherical case are small.

The triangles in Fig. 4 arise from "sloppy" fitness landscapes [34, 35] with the $N$ eigenvalues evenly spaced in the

logarithm from $10^6$ to $10^{-6}$. This corresponds to the narrowest axis of the fitness contours being one-millionth the width of the longest axis. Even for these very non-spherical fitness landscapes, the average spacing $\langle \Delta \rangle$ remains substantial and comparable to the average in the spherical case.

## 4. Discussion

We analyzed adaptive mutation in a version of Fisher's geometric model in which mutations are restricted to acting in only one dimension at a time. This condition of zero pleiotropy is appropriate when the population is described in terms of its chemotype, the biochemical reaction constants of the molecules that comprise the organism, only one or a few of which will be altered by any given point mutation. We showed that the probability density of fitness effects of adaptive mutations will generically exhibit cusps, each associated with mutations of a particular chemotype element. These cusps are particularly prominent in the density of fitness effects of fixed mutations. Simulations suggest that directly resolving these cusps experimentally will be difficult. However, each cusp corresponds to a different element of the chemotype, and we showed that the relative spacing between the two cusps with the highest fitness remains substantial even in very non-spherical landscapes of high dimension. This suggests a testable prediction that is robust to details of the model: the fittest mutations should all affect the same element of the chemotype.

It may be surprising that even very non-spherical fitness landscapes (range in eigenvalues of $10^{12}$) yield a qualitatively similar cusp distribution to the spherical landscape. Our simulations assume that the eigenvectors, and thus the correlations between chemotype elements and fitness, are random. In this case, each chemotype element contributes about equally to each eigenvector, so the fitness function is similar when projected along each chemotype direction, yielding a narrow distribution of $\theta_i$ which is similar to the spherical case. The assumption of random correlation structure is motivated by empirical study of the sensitivity of biochemical networks to reaction constant variation [35] and theoretical study of sloppy systems in general [36]. In both cases, random eigenvectors are a reasonable approximation to the complicated correlations found.

A key assumption of our model is that each chemotype element is continuously adjustable throughout the range of possible adaptive mutations. Because the genetic code is discrete, this cannot be strictly true. The distribution of effects of random mutations on chemotype elements is not well-known, in part because most biochemical studies focus on mutations of large effect. However, studies have shown that random mutations can introduce small but non-zero changes to the enzymatic activity of proteins [37] and the expression driven by promoter sites [38],

suggesting that our assumption of continuous chemotype variation is reasonable.

The substantial average separation $\langle \Delta \rangle$ we predict between the fittest and second-fittest cusp implies that the mutations conveying the largest fitness benefits will all involve a single chemotype element. A similar result holds for the mutational landscape model [15, 39], in which the largest fitness spacings are between the fittest sequences. The spacing distributions in the mutational landscape model, however, depend on the correlation assumed between the effects of different mutations. A recent analysis of such correlations considers mutations within "blocks" of sequence that contribute independently and identically to fitness [40]. Each block may be roughly interpreted as a different chemotype element in our model, but in our model the relative contributions to fitness differ between blocks and naturally arise from the structure of the landscape. Nevertheless, the fact that both models predict the upper end of the fitness distribution to be dominated by few mutations, or in our case mutations in a few chemotype elements, may help explain the large amount of parallel evolution that can be observed in separate populations exposed to similar environments [41, 42].

The distribution of fitness effects of adaptive mutations has been studied in bacteria and viruses [43, 17]. Typically the distribution is found to be consistent with a smooth exponential distribution. Our theory predicts only gentle cusps in this distribution, but much more prominent cusps in the distribution of fitness of effects of fixed mutations. This distribution has been studied experimentally in bacteria [44, 45, 30], and those results are also consistent with a smooth distribution. We showed, however, that it would be difficult for such experiments to directly resolve the cusps. Intriguingly, a recent study of virus adaptation by Rokyta et al. points toward a fitness effects distribution with a truncated right-hand tail [18], consistent with our model.

Directly resolving the cusps is challenging; it will be easier to test the prediction that the fittest mutations will all affect a single element of the chemotype. Given that the average fittest cusp separation $\langle \Delta \rangle$ is roughly 0.1 even for very large $N$, resolving this effect requires a relative precision in fitness of a few percent, which is achievable by averaging repeated assays. Recent developments in microarray-based genotyping [24] allow the sites of mutations to be cost-effectively identified. Mutations that reside in, for example, the same region of a protein likely affect the same element of the chemotype. Correlating fitness measurements of mutations with identification of which chemotype element they affect will allow direct testing of our model predictions.

Motivated by the adaptation of the chemotype, the set of biochemical rate constants comprising an organism, we have studied a version of Fisher's geometrical model without pleiotropy. The model predicts cusps in the distribution of fitness effects of fixed mutations and that the fittest mutations all involve a single element of the chemo-

type. Analysis suggests that the second prediction is experimentally accessible. More broadly, our work suggests that viewing evolution in terms of the chemotype may offer new insights beyond those found at the genotype or phenotype level.

## Acknowledgments

## References

[1] N. Woodford, M. J. Ellington, The emergence of antibiotic resistance by mutation, Clin. Microbiol. Infect. 13 (2007) 5–18.

[2] L. M. F. Merlo, J. W. Pepper, B. J. Reid, C. C. Maley, Cancer as an evolutionary and ecological process, Nat. Rev. Cancer 6 (2006) 924–935.

[3] W. J. Ewens, Mathematical Population Genetics, volume 1, Springer-Verlag, 2nd edition, 2004.

[4] S. F. Elena, R. E. Lenski, Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation, Nat. Rev. Genet. 4 (2003) 457–469.

[5] H. A. Orr, Theories of adaptation: what they do and don't say, Genetica 123 (2005) 3–13.

[6] R. A. Fisher, The genetical theory of natural selection, Oxford Univ Press, Oxford, U.K., 1930.

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

[8] G. Martin, T. Lenormand, The distribution of beneficial and fixed mutation fitness effects close to an optimum, Genetics 179 (2008) 907–916.

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

[10] H. A. Orr, The evolutionary genetics of adaptation: a simulation study, Genet. Res. 74 (1999) 207–214.

[11] D. L. Hartl, C. H. Taubes, Towards a theory of evolutionary adaptation, Genetica 102-103 (1998) 525–533.

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

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

[14] H. A. Orr, The distribution of fitness effects among beneficial mutations in Fisher's geometric model of adaptation, J. Theor. Biol. 238 (2006) 279–285.

[15] H. A. Orr, The population genetics of adaptation: the adaptation of DNA sequences, Evolution 56 (2002) 1317–1330.

[16] D. R. Rokyta, P. Joyce, S. B. Caudle, H. A. Wichman, An empirical test of the mutational landscape model of adaptation using a single-stranded DNA virus, Nat. Genet. 37 (2005) 441–444.

[17] R. Kassen, T. Bataillon, Distribution of fitness effects among beneficial mutations before selection in experimental populations of bacteria, Nat. Genet. 38 (2006) 484–488.

[18] D. R. Rokyta, C. J. Beisel, P. Joyce, M. T. Ferris, C. L. Burch, H. A. Wichman, Beneficial fitness effects are not exponential for two viruses, J. Molec. Evol. 67 (2008) 368–376.

[19] B. C. Daniels, Y.-J. Chen, J. P. Sethna, R. N. Gutenkunst, C. R. Myers, Sloppiness, robustness, and evolvability in systems biology, Curr. Opin. Biotech. 19 (2008) 389 – 395.

[20] D. L. Hartl, D. E. Dykhuizen, A. M. Dean, Limits of adaptation: the evolution of selective neutrality, Genetics 111 (1985) 655–674.

[21] G. Martin, T. Lenormand, A general multivariate extension of Fisher's geometrical model and the distribution of mutation fitness effects across species, Evolution 60 (2006) 893–907.

[22] O. Tenaillon, O. K. Silander, J.-P. Uzan, L. Chao, Quantifying organismal complexity using a population genetic approach, PLoS ONE 2 (2007) e217.

[23] G. P. Wagner, J. P. Kenney-Hunt, M. Pavlicev, J. R. Peck, D. Waxman, J. M. Cheverud, Pleiotropic scaling of gene effects and the 'cost of complexity', Nature 452 (2008) 470–472.

[24] D. Gresham, D. M. Ruderfer, S. C. Pratt, J. Schacherer, M. J. Dunham, D. Botstein, L. Kruglyak, Genome-wide detection of polymorphisms at nucleotide resolution with a single DNA microarray, Science 311 (2006) 1932–1936.

[25] M. Lynch, W. Sung, K. Morris, N. Coffey, C. R. Landry, E. B. Dopman, W. J. Dickinson, K. Okamoto, S. Kulkarni, D. L. Hartl, W. K. Thomas, A genome-wide view of the spectrum of spontaneous mutations in yeast, Proc Natl Acad Sci U S A 105 (2008) 9272–9277.

[26] J. R. Peck, G. Barreau, S. C. Heath, Imperfect genes, Fisherian mutation and the evolution of sex, Genetics 145 (1997) 1171–1199.

[27] D. Waxman, J. J. Welch, Fisher's microscope and Haldane's ellipse, Am. Nat. 166 (2005) 447–457.

[28] G. Martin, T. Lenormand, The fitness effect of mutations across environments: a survey in light of fitness landscape models, Evolution 60 (2006) 2413–2427.

[29] A. Eyre-Walker, P. D. Keightley, The distribution of fitness effects of new mutations, Nat. Rev. Genet. 8 (2007) 610–618.

[30] L. Perfeito, L. Fernandes, C. Mota, I. Gordo, Adaptive mutations in bacteria: high rate and small effects, Science 317 (2007) 813–815.

[31] G. Sella, A. E. Hirsh, The application of statistical physics to evolutionary biology, Proc. Natl. Acad. Sci. USA 102 (2005) 9541–9546.

[32] E. J. Gumbel, Statistics of Extremes, Columbia University Press, New York, 1958.

[33] D. Waxman, Mean curvature versus normality: a comparison of two approximations of Fisher's geometrical model, Theor. Popul. Biol. 71 (2007) 30–36.

[34] K. S. Brown, J. P. Sethna, Statistical mechanical approaches to models with many poorly known parameters, Phys. Rev. E 68 (2003) 021904.

[35] R. N. Gutenkunst, J. J. Waterfall, F. P. Casey, K. S. Brown, C. R. Myers, J. P. Sethna, Universally sloppy parameter sensitivities in systems biology, PLoS Comput. Biol. 3 (2007) e189.

[36] J. J. Waterfall, F. P. Casey, R. N. Gutenkunst, K. S. Brown, C. R. Myers, P. W. Brouwer, V. Elser, J. P. Sethna, Sloppy-model universality class and the Vandermonde matrix, Phys. Rev. Lett. 97 (2006) 150601.

[37] J. D. Bloom, P. A. Romero, Z. Lu, F. H. Arnold, Neutral genetic drift can alter promiscuous protein functions, potentially aiding functional evolution, Biol. Direct 2 (2007) 17.

[38] C. R. Landry, B. Lemos, S. A. Rifkin, W. J. Dickinson, D. L. Hartl, Genetic properties influencing the evolvability of gene expression, Science 317 (2007) 118–121.

[39] H. A. Orr, A minimum on the mean number of steps taken in adaptive walks, J. Theor. Biol. 220 (2003) 241–247.

[40] H. A. Orr, The population genetics of adaptation on correlated fitness landscapes: the block model, Evolution 60 (2006) 1113–1124.

[41] T. E. Wood, J. M. Burke, L. H. Rieseberg, Parallel genotypic adaptation: when evolution repeats itself, Genetica 123 (2005) 157–170.

[42] L. Pelosi, L. Khn, D. Guetta, J. Garin, J. Geiselmann, R. E. Lenski, D. Schneider, Parallel changes in global protein profiles during long-term experimental evolution in Escherichia coli, Genetics 173 (2006) 1851–1869.

[43] M. Imhof, C. Schlotterer, Fitness effects of advantageous mutations in evolving escherichia coli populations, Proc. Natl. Acad. Sci. USA 98 (2001) 1113–1117.

[44] D. E. Rozen, J. A. G. M. de Visser, P. J. Gerrish, Fitness effects of fixed beneficial mutations in microbial populations, Curr. Biol. 12 (2002) 1040–1045.

[45] R. D. H. Barrett, R. C. MacLean, G. Bell, Mutations of intermediate effect are responsible for adaptation in evolving Pseudomonas fluorescens populations, Biol. Lett. 2 (2006) 236–238.

[46] M. Abramowitz, I. A. Stegun, Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, Dover, New York, 10th edition, 1964.

## Appendix A. Scale of chemotype mutation

If the distribution $f(r)$ of mutation effects on the chemotype were small for $r$ greater than the typical $\rho_i$, the mutational distance over which adaptive mutations are possible, then the fraction $P_a$ of mutations that were adaptive would be approximately one-half. The rarity of adaptive mutations thus suggests that $f(r)$ must be appreciable for $r$ greater than the typical $\rho_i$. We now show that $f(r)$ must remain substantial even for $r$ greater than the *largest* $\rho_i$. To do so, we make the simplifying assumption that the distribution of mutational effects is identical for all chemotype elements. We then consider the scenario in which this distribution barely covers the range of all possible adaptive mutations, extending only to the largest of the $\rho_i$, demoted $\max_i \rho_i$. For this scenario, we derive an analytic approximation to $P_a$ for spherical landscapes, and we calculate $P_a$ numerically for the non-spherical landscapes considered in Fig. 4. In both cases we find that this scenario leads to an unrealistically high probability of adaptive mutation, implying that the distribution of mutation chemotype effects must have a scale larger than that of the largest possible adaptive mutation.

If the probability density of mutation chemotype effects $f(r)$ were uniform over $(-\max_i \rho_i, +\max_i \rho_i)$, the probability of a random mutation being adaptive would be

$$P_a = \frac{\sum_i \rho_i}{2N \max_i \rho_i}. \tag{A.1}$$

The numerator is the total length of intervals where mutations are adaptive, and the denominator is the total length of intervals over which mutations are distributed.

Specializing to the spherical case and plugging in for $\rho_i$, we have

$$P_a = \frac{\left\langle |\hat{k}\cdot\hat{r}_i| \right\rangle}{2 \max_i |\hat{k}\cdot\hat{r}_i|}. \tag{A.2}$$

Asymptotically for large $N$, $\hat{k}\hat{r}_i$ has a Gaussian probability density with variance $1/N$. Averaging yields

$$\left\langle |\hat{k}\cdot\hat{r}_i| \right\rangle = \sqrt{2}/\sqrt{\pi N}. \tag{A.3}$$

The largest absolute value of $N$ samples drawn from a Gaussian density with variance $1/N$ is asymptotically $\sqrt{2 \log \left( N/\sqrt{2\pi} \right)/N}$ [32, Eq. 4.2.3(11)]. Plugging these into Eq. A.2 yields

$$P_a(N) \sim \frac{1}{2\sqrt{\pi \log(N/\sqrt{2\pi})}}. \tag{A.4}$$

This probability remains substantial even for very large $N$. For example, $P_a(10,000)$ is roughly 0.1, an unrealistically large value.

Numerical tests with both the mildly and wildly non-spherical landscapes considered in Fig. 4 yield a $P_a$ of at least 0.13, consistent with the spherical result of Eq. A.4. This suggests that, for a realistic fraction of mutations to be deleterious, the typical scale of chemotype effects for mutations must be larger than $\max_i \rho_i$, even for very non-spherical landscapes.

## Appendix B. Fitness effects distribution

Given the probability density of chemotype mutation effects $f(r)$, the distribution $f_a(Q)$ of fitness effects for adaptive mutations is

$$f_a(Q) \propto \sum_i \int dr f(r) \, \delta(Q - Q_i(r)), \tag{B.1}$$

where the sum is over all chemotype elements $i$, and $Q_i(r)$ is given by Eq. 3. Making the variable substitution $u = Q_i(r)$ yields

$$f_a(Q) \propto \sum_i \int du \frac{f(Q_i^{-1}(u)) \, \delta(Q - u)}{\sqrt{\left| \vec{k} \cdot \mathbf{S} \cdot \hat{r}_i \right|^2 - 2 \, \hat{r}_i \cdot \mathbf{S} \cdot \hat{r}_i \, u}}. \tag{B.2}$$

Further making the approximation that $f(r)$ is uniform over the range of adaptive mutations, and substituting Eq. 4 yields Eq. 8.

## Appendix C. Average distributions

The distribution of adaptive fitness effects averaged over initial chemotypes with a given fitness can be calculated by averaging $f_a(Q)$ over the probability density for $\theta_i$. For a spherical fitness landscape, the $\theta_i$ are proportional to the squared magnitudes of the components of the unit vector $\hat{k}$. Asymptotically as the number of dimensions $N \to \infty$, these are squares of Gaussian variables and have probability density

$$f_\theta(\theta_i) \propto \exp\left[-\theta_i N/(2Q_{tot})\right]/\sqrt{\theta_i}, \tag{C.1}$$

which is a $\chi^2$ density with one degree of freedom. For the spherical fitness landscape the result is:

$$f_{a,e}(Q) \propto \exp\left[-QN/(4Q_{tot})\right] K_0\left[QN/(4Q_{tot})\right], \tag{C.2}$$

where $K_0$ is the zero-order modified Bessel function of the second kind. This is identical to Waxman and Welch's result for the model with maximum pleiotropy [27].

## Appendix D. Extreme value theory for $\Delta$

The normalized spacing between cusps $\Delta$ is a ratio of two values; to calculate its probability density we first calculate the density of $i_1 \equiv \log \theta_1 - \log \theta_2$, the spacing between the logarithms of the largest two $\theta$s. Defining

$$\omega \equiv \log\left(\frac{\theta N}{Q_{tot}}\right) \tag{D.1}$$

and using the asymptotic $\chi^2$ density for $\theta$ yields the asymptotic probability density of $\omega$:

$$f(\omega) = \exp\left[-\frac{1}{2}\left(\exp(\omega) - \omega\right)\right]/\sqrt{2\pi}. \tag{D.2}$$

The corresponding cumulative probability distribution $F(\omega) \equiv \int_{-\infty}^{\omega} f(\omega')d\omega'$ is

$$F(\omega) = \mathrm{erf}\left(\exp(\omega/2)/\sqrt{2}\right), \tag{D.3}$$

where erf is the error function. This distribution has exponential-type extreme value statistics [32].

Following the terminology and notation of Gumbel [32], the typical size $u_{1,N}$ of the largest of N samples from the density $f(\omega)$ is given by $F(u_{1,N}) = 1 - \frac{1}{N}$. In our case this is

$$u_{1,N} = 2 \log\left(\sqrt{2}\,\mathrm{erf}^{-1}\left(1 - 1/N\right)\right). \tag{D.4}$$

The corresponding scale parameter $\alpha_{1,N}$ is

$$\alpha_{1,N} = N f\left(u_{1,N}\right), \tag{D.5}$$

and distance $i_1$ between the largest two samples has probability density

$$f(i_1) = \alpha_{1,N} \exp(-\alpha_{1,N}\, i_1). \tag{D.6}$$

(Gumbel's result [32] for this distribution, his Eq. 5.3.5(4), has $\alpha_{2,N}$ in place of $\alpha_{1,N}$. In the limit $N \to \infty$ the two expressions are equal, but $\alpha_{1,N}$ is a better approximation for small $N$.)

The distance between the logarithms $i_1$ is related to $\Delta$ by $\Delta \equiv 1 - \theta_2/\theta_1 = 1 - \exp(-i_1)$. Thus the probability density for $\Delta$ is

$$f(\Delta) = \alpha_{1,N}\left(1 - \Delta\right)^{(\alpha_{1,N} - 1)}, \tag{D.7}$$

and the average of $\Delta$ is

$$\langle \Delta \rangle = \frac{1}{1 + \alpha_{1,N}}. \tag{D.8}$$

A useful approximation for $\alpha_{1,N}$ can be obtained using an asymptotic expansion for $\mathrm{erf}^{-1}$ [46]:

$$\sqrt{2}\,\mathrm{erf}^{-1}\left(1 - x\right) \sim \sqrt{\log\left(\frac{2}{\pi x^2}\right) - \log\log\left(\frac{2}{\pi x^2}\right)}. \tag{D.9}$$

Propagating this expansion through $u_{1,N}$ and $\alpha_{1,N}$ and neglecting terms of order $\log \log N$ in the final expression yields

$$\alpha_{1,N} \approx \log N + \log\left(\sqrt{2/\pi}\right). \qquad \text{(D.10)}$$

From this follows the approximate expression for $\langle \Delta \rangle$ in Eq. 11.

## Appendix E. Numerical simulation

A random set of orthogonal unit vectors $\hat{v}_i$ can be obtained from the eigenvectors of a matrix $\mathbf{G}$ from the Gaussian Orthogonal Ensemble: $\mathbf{G} = \mathbf{H} + \mathbf{H}^{\mathsf{T}}$ where the elements of $\mathbf{H}$ are standard normal random numbers. A matrix $\mathbf{S}$ with eigenvalues $\lambda_i$ can then be constructed via

$$S_{j,k} = \sum_i \lambda_i v_{i,j} v_{i,k}. \qquad \text{(E.1)}$$

Random chemotypes $\vec{k}$ with specified log-fitness $Q_{tot} = -\log W(\vec{k})$ are obtained using the Cholesky decomposition $\mathbf{A}$ of $\mathbf{S}^{-1}$, defined by $\mathbf{A} \cdot \mathbf{A}^{\mathsf{T}} = \mathbf{S}^{-1}$. $\vec{k}$ is then given by

$$\vec{k} = \sqrt{2Q_{tot}}\mathbf{A} \cdot \hat{k}, \qquad \text{(E.2)}$$

where $\hat{k}$ is a random unit vector.