

14. S. J. Simpson, C. L. Simpson, *J. Exp. Biol.* **168**, 269 (1992).
15. J. D. Abisgold, S. J. Simpson, *J. Exp. Biol.* **135**, 215 (1988).
16. C. L. Simpson, S. Chyb, S. J. Simpson, *Entomol. Exp. Appl.* **56**, 259 (1990).
17. R. F. Chapman, *Annu. Rev. Entomol.* **48**, 455 (2003).
18. R. M. Sibly, in *Physiological Ecology: An Evolutionary Approach to Resource Use*, T. C. P. Calow, Ed. (Blackwell, Oxford, 1981), pp. 109–139.
19. L. A. Real, *Am. Nat.* **140**, S108 (1992).
20. E. A. Bernays, *Annu. Rev. Entomol.* **46**, 703 (2001).
21. G. Gigerenzer, *Behav. Brain Sci.* **27**, 336 (2004).
22. We thank C. Nadell and C. Spatola Rossi for help during data collection; C. Schuck-Paim and M. S. Shapiro for helpful discussion; and T. Dickinson, P. Harvey, T. Reil, O. Lewis, and two anonymous referees for valuable feedback on the manuscript.

Supporting Online Material

www.sciencemag.org/cgi/content/full/311/5767/1613/DC1

Materials and Methods

Figs. S1 and S2

References

15 December 2005; accepted 1 February 2006

10.1126/science.1123924

An Equivalence Principle for the Incorporation of Favorable Mutations in Asexual Populations

Matthew Hegreness,^{1,2*} Noam Shores^{1,*} Daniel Hartl,² Roy Kishony^{1,3†}

Rapid evolution of asexual populations, such as that of cancer cells or of microorganisms developing drug resistance, can include the simultaneous spread of distinct beneficial mutations. We demonstrate that evolution in such cases is driven by the fitness effects and appearance times of only a small minority of favorable mutations. The complexity of the mutation-selection process is **thereby greatly reduced, and much of the evolutionary dynamics can be encapsulated in two parameters—an effective selection coefficient and effective rate of beneficial mutations.** We confirm this theoretical finding and estimate the effective parameters for evolving populations of fluorescently labeled *Escherichia coli*. The effective parameters constitute a simple description and provide a natural standard for comparing adaptation between species and across environments.

Spontaneous beneficial mutations are the fuel for adaptation, the source of evolutionary novelty, and one of the least understood aspects of biology. Although adaptation is everywhere—cancer invading tissues, bacteria escaping drugs, viruses switching from livestock to humans—beneficial mutations are notoriously difficult to study (1, 2). **Theoretical and experimental advances have been made in recent years by focusing on the distribution of fitness effects of spontaneous beneficial mutations (3–8).** Mapping the options for improvement available to single organisms, however, is insufficient for understanding the adaptive course of an entire population, especially in asexual populations of microorganisms or cancer cells where multiple mutations often spread simultaneously (9–16). Here, we use modeling and experimental results to show that the seeming additional complication of having multiple lineages competing within a population leads in fact to a drastic simplification: **Regardless of the distribution of mutational effects available to individuals, a population's adaptive dynamics can be approximated by an equivalent model in which all favorable mutations confer the same fitness advantage, which we call the effective selection coefficient. We provide experimental estimates of the effective selection coefficient**

and the corresponding effective rate of beneficial mutations for laboratory populations of *Escherichia coli*, and we demonstrate the predictive power of these effective parameters.

First, we use numerical simulations to demonstrate the simplification that emerges in a population large enough and a mutation rate high enough that clonal interference (17–19)—competition among lineages carrying favorable mutations—is common. In an evolving population, most beneficial mutations are rapidly lost to random genetic drift (20, 21). Of the remaining mutant lineages, some increase in frequency slightly, only to decline as more fit lineages appear and expand in the population (10, 16, 17, 22). The evolutionary path taken by the population as a whole is determined by successful mutations that escape stochastic loss and whose frequencies rise above some minimal level. Using a population genetics model that includes mutation, selection, drift, as well as clonal interference (23), we explore the distribution of these successful mutations for several underlying distributions of

beneficial mutations (Fig. 1), including an exponential distribution as suggested by Gillespie's (8) and Orr's (3) use of extreme value theory.

The salient feature of Fig. 1 is that very dissimilar underlying distributions—exponential, uniform, lognormal, even an arbitrary distribution—all yield a similar distribution of successful mutations (24). Moreover, the distribution of successful mutations has a simple form, peaked around a single value. This fitness value is typical of those mutations whose effects are not so small that they are lost through competition with more fit lineages, but are also not so large that they are impossibly rare. The unimodal shape motivates the hypothesis that an equivalent model that allows mutations with only a single selective value might approximate the behavior of the entire distribution of beneficial mutations.

We investigate whether the adaptive dynamics observed in evolving *E. coli* populations can be reproduced by an equivalent model with only a single value, a Dirac delta function of mutational effects. We rely on a classic strategy for characterizing beneficial mutations in coevolving subpopulations that differ initially only by selectively neutral marker. The spread of mutations is monitored through changes in the marker ratio (22, 25–29). Our experimental technique uses constitutively expressed variants of GFP (green fluorescent protein)—YFP (yellow fluorescent protein) and CFP (cyan fluorescent protein)—as neutral markers. All experimental populations start with equal numbers of YFP and CFP *E. coli* cells (N_Y and N_C) and evolve for 300 generations through serial transfers while adapting to glucose minimal medium.

The expected behavior of the marker-ratio trajectories depends upon the rate at which beneficial mutations appear in a population. When beneficial mutations are rare, mutant lineages arise and fix one at a time (8, 17). The spread of each individual mutant lineage shows as a line of

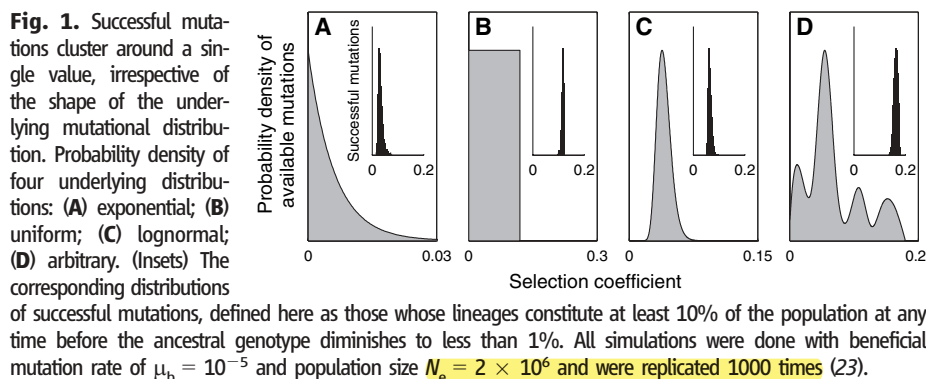


Fig. 1. Successful mutations cluster around a single value, irrespective of the shape of the underlying mutational distribution. Probability density of four underlying distributions: (A) exponential; (B) uniform; (C) lognormal; (D) arbitrary. (Insets) The corresponding distributions of successful mutations, defined here as those whose lineages constitute at least 10% of the population at any time before the ancestral genotype diminishes to less than 1%. All simulations were done with beneficial mutation rate of $\mu_b = 10^{-5}$ and population size $N_e = 2 \times 10^6$ and were replicated 1000 times (23).

¹Bauer Center for Genomics Research, Harvard University, Cambridge, MA 02138, USA. ²Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA. ³Department of Systems Biology, Harvard Medical School, Boston, MA 02115, USA.

*These authors contributed equally to this work.

†To whom correspondence should be addressed. E-mail: roy_kishony@hms.harvard.edu

Fig. 2. The evolutionary spread of beneficial mutations. (A) A mutation (dark yellow) that occurs in a YFP-labeled cell takes over a mixed population of YFP and CFP cells. (B) The observed N_Y/N_C ratio in the population from (A); the mutant's selection coefficient, s_Y , can be obtained from the slope at late times. (C) A beneficial mutation in YFP competing with a beneficial mutation in CFP that occurs later but has a stronger selective advantage, s_C . (D) N_Y/N_C in the population from (C); the slope $d\log_2(N_Y/N_C)/dt$ at late times is equal to $s_Y - s_C$.

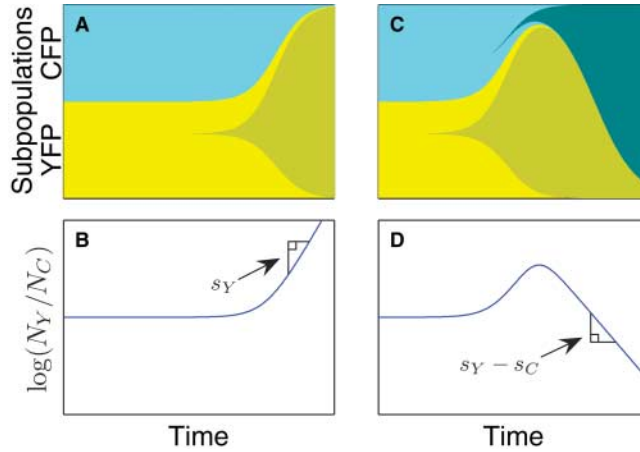


Fig. 3. Empirical and numerical marker-ratio data and trajectory parameterization. (A) Ratio of YFP to CFP cells of *E. coli* monitored for 300 generations (for clarity, 36 of 72 populations are shown and the rest are presented in fig. S2). The plateaus and reversals (see examples in bold) reveal the simultaneous spread of distinct beneficial mutations. The asymmetry in the vertical axis reflects the asymmetry in the dynamic range due to the fluorescence properties of YFP and CFP. (B) Ratios based on simulations using the Dirac delta function with a beneficial mutation rate $\mu_e = 10^{-6.7}$ and a selection coefficient $s_e = 0.054$ (values obtained from region of agreement in Fig. 4). (C) α and τ are defined by the function shown. The data, the best fit curve, and the corresponding values for α and τ are shown from one experimental population. (D and E) Histograms of estimated values of α and τ for each of the 72 experimental populations (for the correlation between α and τ see fig. S5).

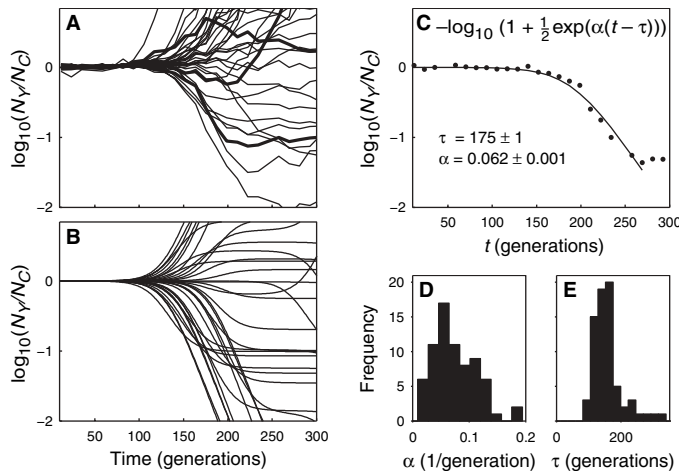
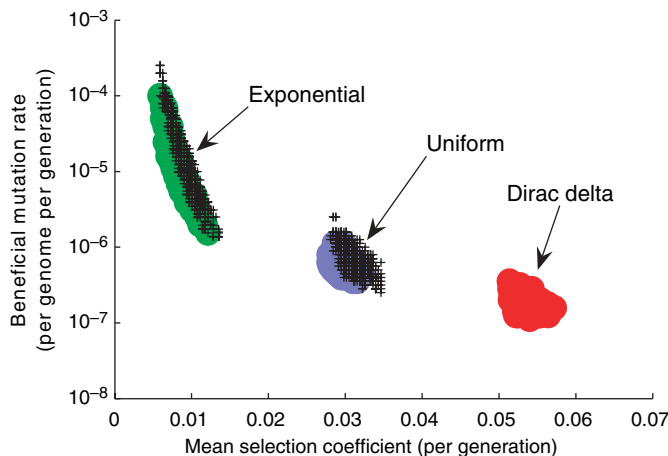


Fig. 4. Estimation of the effective parameters and demonstration that their predictive power is robust to the underlying distribution. Filled regions represent agreement (at 95% confidence level) between empirical data and model simulations, based on comparison of marker-ratio trajectories; shown in red is the region of agreement for the Dirac delta function, which provides an estimate of the effective parameters; also shown are the regions of agreement for exponential (green) and uniform (blue) underlying distributions. Crosshatched areas are the regions where calculations of $\langle f_{\max} \rangle$ and $\langle s_{\text{succ}} \rangle$ from the exponential and uniform distributions agree with predictions of these quantities derived from the effective parameters.



constant nonzero slope when the logarithm of the marker ratio is plotted against time (Fig. 2A), where the slope is equal to the selection coefficient of the expanding lineage (27). When the mutation rate in asexual populations is high, however, beneficial mutations arise in both subpopulations and compete (Fig. 2B).

Results of the adaptation experiments are shown in Fig. 3A. As expected, the curves of the logarithm of the ratio of N_Y to N_C are initially flat, reflecting the equal initial fitness of the ancestral YFP and CFP cells. Beneficial mutations cause the marker ratio to deviate from its starting value after ~ 100 generations. The plateaus and reversals of the slopes that often appear after these initial deviations reveal the simultaneous spread of multiple beneficial mutations (additional evidence for the presence of clonal interference is shown in figs. S3 and S4). Concentrating on the initial phase of the experiment, we extract the time, τ , when significant deviation from a flat line first occurs, and the slope, α , of that deviation (Fig. 3C). We extract these for all 72 marker-ratio trajectories, obtaining empirical samples of α and τ (Fig. 3, D and E).

Results from the evolving *E. coli* populations are compared to simulations results produced with a theoretical model. The model tracks two coevolving subpopulations (“cyan” and “yellow”) and accounts for the fluctuating population size due to serial dilution, selection, and random drift (23). An input to the model is a distribution from which the selection coefficients of beneficial mutations are drawn.

To test whether mutations of a single effect can generate variable adaptive dynamics compatible with the empirical data, we use a Dirac delta function as the equivalent underlying distribution of beneficial mutations. We also explore two other one-parameter distribution families—uniform and exponential (to facilitate comparison, distributions with more than one parameter, including the log-normal distribution from Fig. 1, are excluded from this analysis, though the conclusions that follow apply to these as well). Given a mutational distribution, the model has only two free parameters: the beneficial mutation rate and the mean of the distribution. For each distribution, for each point in parameter space, many realizations of the model are simulated (see example in Fig. 3B), generating theoretical predictions in the form of numerical samples of α and τ . These numerical samples, produced with the model, are then compared to the corresponding empirical ones from the *E. coli* experiments (using a Kolmogorov-Smirnov test). The filled areas in Fig. 4 indicate the region of agreement between the model and the empirical data for each of the underlying distributions of beneficial mutations.

The agreement with the Dirac delta function demonstrates that beneficial mutations of a single magnitude can indeed give rise to the rich adaptive behavior observed in the experiment. In particular, the differences in the timings of the mutations are a sufficient source for the variability

ity in the marker-ratio data. The effective parameters of the bacterial populations are estimated from the delta function's region of agreement: the effective selection coefficient $s_e = 0.054 \pm 0.003$ and the effective rate $\mu_e = 10^{-6.7 \pm 0.2}$ mutations per genome per generation.

The consistency of more than one underlying distribution with the data reinforces the point illustrated in Fig. 1: Adaptive dynamics are largely determined by a few broad properties of the distribution, encapsulated by the effective parameters, and not by its exact shape. The regions of agreement, obtained by comparing marker-ratio trajectories, are thus interpreted as reducing to the same effective parameters.

The equivalent model (defined by the effective parameters) predicts other measures of adaptation not trivially related to the marker-ratio data. Using simulations, we examine two quantities: $\langle f_{\max} \rangle$, the degree of polymorphism, as measured by the average (over repeated simulations) of the frequency of the most common beneficial mutation when the ancestral strain goes extinct (reduces to 1% of the population); and $\langle s_{\text{suc}} \rangle$, the mean effect of successful mutations, that is, the average fitness of all mutant lineages that ever reach more than 10% frequency by this time. Predictions of these quantities based on the effective parameters are obtained through model simulations with the Dirac delta function. Regions where calculations of $\langle f_{\max} \rangle$ and $\langle s_{\text{suc}} \rangle$ from the exponential and uniform distributions agree with these predictions are shown as the crosshatched areas in Fig. 4. The striking overlap of the new regions with the regions estimated from the marker-ratio data highlights the diversity of features captured by the effective parameters. Comparing the Dirac delta function with a more complicated underlying distribution, we see that if the distributions agree in their predictions for one aspect of adaptation (marker-ratio trajectories), they will also agree in other aspects ($\langle f_{\max} \rangle$ and $\langle s_{\text{suc}} \rangle$). In accordance with the idea of an equivalent model, these results suggest that the predictive potential of the effective parameters is independent of the actual underlying mutational distribution.

When multiple beneficial mutations spread simultaneously in asexual populations, adaptive dynamics can be reasonably described by an equivalent model in which all favorable mutations confer the same selection advantage. The scope of this simple approximation, namely, the breadth of observables it captures and the limits at which it breaks down, is the subject of future research. Like a local adaptive landscape, the effective selection coefficient and the effective rate of beneficial mutations characterize the dynamics of a population at a specific point in its evolution. An entire adaptive trajectory might be represented by tracking changes in the effective parameters. Compared to high-dimensional fitness landscapes, effective parameters constitute a major simplification and can serve as mileposts along the adaptive walk.

References and Notes

1. J. J. Bull, S. P. Otto, *Nat. Genet.* **37**, 342 (2005).
2. H. A. Orr, *Evolution* **52**, 935 (1998).
3. H. A. Orr, *Genetics* **163**, 1519 (2003).
4. D. R. Rokyta, P. Joyce, S. B. Caudle, H. A. Wichman, *Nat. Genet.* **37**, 441 (2005).
5. R. Sanjuan, A. Moya, S. F. Elena, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 8396 (2004).
6. M. A. DePristo, D. M. Weinreich, D. L. Hartl, *Nat. Rev. Genet.* **6**, 678 (2005).
7. M. Lunzer, S. P. Miller, R. Felsheim, A. M. Dean, *Science* **310**, 499 (2005).
8. J. H. Gillespie, *The Causes of Molecular Evolution* (Oxford Univ. Press, New York, 1991).
9. R. E. Lenski, M. R. Rose, S. C. Simpson, S. C. Tadler, *Am. Nat.* **138**, 1315 (1991).
10. R. Miralles, P. J. Gerrish, A. Moya, S. F. Elena, *Science* **285**, 1745 (1999).
11. J. A. G. M. de Visser, C. W. Zeyl, P. J. Gerrish, J. L. Blanchard, R. E. Lenski, *Science* **283**, 404 (1999).
12. N. Colegrave, *Nature* **420**, 664 (2002).
13. A. Moya, E. C. Holmes, F. Gonzalez-Candelas, *Nat. Rev. Microbiol.* **2**, 279 (2004).
14. G. H. Heppner, *Cancer Res.* **44**, 2259 (1984).
15. M. R. Novelli *et al.*, *Science* **272**, 1187 (1996).
16. J. Kim, L. R. Ginzburg, D. E. Dykhuizen, in *Assessing Ecological Risks of Biotechnology*, L. R. Ginzburg, Ed. (Butterworth-Heinemann, Boston, 1991), pp. 193–214.
17. P. J. Gerrish, R. E. Lenski, *Genetica* **102-103**, 127 (1998).
18. H. J. Muller, *Am. Nat.* **66**, 118 (1932).
19. J. F. Crow, M. Kimura, *Am. Nat.* **99**, 439 (1965).
20. M. Kimura, *The Neutral Theory of Molecular Evolution* (Cambridge Univ. Press, New York, 1983).
21. J. Haldane, *Proc. Cambridge Philos. Soc.* **23**, 838 (1927).
22. D. E. Rozen, J. de Visser, P. J. Gerrish, *Curr. Biol.* **12**, 1040 (2002).
23. Materials and methods are available as supporting material on Science Online.
24. A similar observation was made by Gerrish and Lenski in (17) concerning the probability of transiently polymorphic beneficial mutations.
25. K. C. Atwood, L. K. Schneider, F. J. Ryan, *Proc. Natl. Acad. Sci. U.S.A.* **37**, 146 (1951).
26. D. L. Hartl, D. E. Dykhuizen, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 6344 (1981).
27. L. Chao, E. C. Cox, *Evolution* **37**, 125 (1983).
28. M. M. Zambrano, D. A. Siegle, M. Almiron, A. Tormo, R. Kolter, *Science* **259**, 1757 (1993).
29. M. Imhof, C. Schlotterer, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 1113 (2001).
30. We thank D. Huse for theoretical advice regarding the model; M. Elowitz for strains and suggestions; N. Barkai and R. Lenski for important comments; and U. Alon, D. Damian, A. DeLuna, M. DePristo, M. Desai, D. Fisher, L. Garwin, A. Murray, V. Savage, K. Vestigian, and D. Weinreich for discussions and comments on the manuscript. This work was supported by the Bauer Center for Genomics Research and by grants from the NSF and the Ellison Medical Foundation. M.H. dedicates this work to the memory of his mother, Kari Hegreness.

Supporting Online Material

www.sciencemag.org/cgi/content/full/311/5767/1615/DC1
Materials and Methods

SOM Text

Figs. S1 to S5

References

10 November 2005; accepted 9 February 2006

10.1126/science.1122469

Parietal-Eye Phototransduction Components and Their Potential Evolutionary Implications

Chih-Ying Su,^{1,*†} Dong-Gen Luo,¹ Akihisa Terakita,² Yoshinori Shichida,² Hsi-Wen Liao,¹ Manija A. Kazmi,³ Thomas P. Sakmar,³ King-Wai Yau^{1*}

The parietal-eye photoreceptor is unique because it has two antagonistic light signaling pathways in the same cell—a hyperpolarizing pathway maximally sensitive to blue light and a depolarizing pathway maximally sensitive to green light. Here, we report the molecular components of these two pathways. We found two opsins in the same cell: the blue-sensitive pinopsin and a previously unidentified green-sensitive opsin, which we name parietopsin. Signaling components included gustducin- α and G_{α_i} , but not rod or cone transducin- α . Single-cell recordings demonstrated that G_o mediates the depolarizing response. Gustducin- α resembles transducin- α functionally and likely mediates the hyperpolarizing response. The parietopsin- G_o signaling pair provides clues about how rod and cone phototransduction might have evolved.

Lizards and some other lower vertebrates have a third eye (parietal eye) (*1*) in addition to the two lateral eyes. This eye may mediate the global detection of dawn and dusk (*1, 2*) instead of conventional image-forming vision. The parietal-eye photoreceptors resemble rods and cones in morphology, but they show chromatic antagonism (a unique feature among all known photoreceptors) consisting of a hyperpolarizing light response most sensitive to blue light and a depolarizing light response most sensitive to green light (*2*). The hyperpolarizing response is produced, as in rods and cones, by the activation of a gua-

nosine 3',5'-cyclic monophosphate (cGMP)-phosphodiesterase that lowers the cGMP concentration and closes cyclic nucleotide-gated

¹Department of Neuroscience, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA. ²Department of Biophysics, Graduate School of Science, Kyoto University and Core Research for Evolutional Science and Technology, Japan Science and Technology Agency, Kyoto 606-8502, Japan. ³Laboratory of Molecular Biology and Biochemistry, Rockefeller University, 1230 York Avenue, New York, NY 10021, USA.

*To whom correspondence should be addressed. E-mail: chih-ying.su@yale.edu (C.-Y.S.); kwau@mail.jhmi.edu (K.-W.Y.)

†Present address: Department of Molecular, Cellular, and Developmental Biology, Yale University, New Haven, CT 06520, USA.