united into a single island, from which the Bismarcks were always separate. A plausible interpretation of the Papuan language tree is thus that the two language groups now located on the Solomons and Bougainville separated from a common ancestor. This could have happened while they could still freely migrate on a common landmass, a time depth ($\sim$10,000 years) in accord with that required to erode traces of common vocabulary. This population history hypothesis will require further testing with both linguistic and genetic data.

If grammatical structures can retain a phylogenetic signal beyond the current temporal ceiling on the reconstruction of language history, then the possibility is opened up of finding relationships between others of the world's 300 or so existing language families and isolates.

References and Notes

1. J. Nichols, *Linguistic Diversity in Space and Time* (Univ. of Chicago Press, Chicago, 1992).
2. J. Nichols, in *The Handbook of Historical Linguistics*, B. D. Joseph, R. D. Janda, Eds. (Blackwell, Oxford, 2003), pp. 283–310.
3. J. H. Greenberg, *Language in the Americas* (Stanford Univ. Press, Stanford, CA, 1987).
4. L. L. Cavalli-Sforza, P. Menozzi, A. Piazza, *The History and Geography of Human Genes* (Princeton Univ. Press, Princeton, NJ, 1994).
5. D. Bolnick, B. Shook, L. Campbell, I. Goddard, *Am. J. Hum. Genet.* **75**, 519 (2004).
6. P. Kirch, *The Lapita Peoples* (Blackwell, London, 1997).
7. M. Spriggs, *The Island Melanesians* (Blackwell, London, 1997).
8. D. A. Merriwether *et al.*, in *Genomic Diversity: Applications in Human Population Genetics*, S. S. Papiha, R. Deka, R. Chakraborty, Eds. (Kluwer Academic/Plenum, New York, 1999), p. 153.
9. M. Kayser *et al.*, *Am. J. Hum. Genet.* **72**, 281 (2003).
10. J. Lynch, M. Ross, T. Crowley, *The Oceanic Languages* (Curzon Press, Richmond, UK, 2002).
11. M. Ross, in *The Boy from Bundaberg: Studies in Melanesian Linguistics in Honour of Tom Dutton*, A. Pawley, M. Ross, D. Tryon, Eds. (Pacific Linguistics, Canberra, Australia, 2001), pp. 301–321.
12. S. A. Wurm, *Papuan Languages of Oceania* (Gunter Narr Verlag, Tübingen, Germany, 1982).
13. J. H. Greenberg, in *Current Trends in Linguistics*, vol. 8, *Linguistics in Oceania*, T. A. Sebeok, Ed. (Mouton and Co., the Hague, 1971), pp. 807–871.
14. D. Ringe, T. Warnow, A. Taylor, *Trans. Philol. Soc.* **100**, 59 (2002).
15. P. Forster, A. Toth, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 9079 (2003).
16. R. D. Gray, Q. D. Atkinson, *Nature* **426**, 435 (2003).
17. R. D. Gray, F. M. Jordan, *Nature* **405**, 1052 (2000).
18. C. J. Holden, R. Mace, *Proc. R. Soc. London Ser. B* **270**, 2425 (2003).
19. C. J. Holden, *Proc. R. Soc. London Ser. B* **269**, 793 (2001).
20. K. Rexová, D. Frynta, D. J. Zrzavý, *Cladistics* **19**, 120 (2003).
21. A. McMahon, R. McMahon, *Trans. Philol. Soc.* **101**, 7 (2003).
22. J. Nichols, in *Historical Linguistics 1993. Papers from the Eleventh International Conference on Historical Linguistics*, H. Andersen, Ed. (John Benjamins, Amsterdam, 1995), pp. 337–356.
23. J. Nichols, in *The Origin and Diversification of Language*, N. Jablonski, L. C. Aiello, Eds. (California Academy of Sciences, San Francisco, 1998), pp. 127–170.
24. M. Dunn, G. Reesink, A. Terrill, *Ocean. Linguist.* **41**, 28 (2002).
25. See sources of language data in the supporting online material.
26. Materials and methods and a description of linguistic characters are available as supporting material on *Science* Online.
27. M. Ross, *Proto Oceanic and the Austronesian Languages of Western Melanesia* (Pacific Linguistics C-98, Canberra, Australia, 1988).
28. This intermeshing of trees does reflect long-term contact in New Britain (*26*).
29. D. Tryon, B. Hackman, *Solomon Island Languages: An Internal Classification* (Pacific Linguistics C-72, Canberra, Australia, 1983).
30. This work, as part of the European Science Foundation EUROCORES Programme OMLL, was supported by funds from the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO); the Arts and Humanities Research Council, UK; and the EC Sixth Framework Programme under contract no. ERAS-CT-2003-980409. Additional fieldwork data used in this study were provided by E. Lindström, S. Robinson, T. Stebbins, W. Thurston, and C. Wegener; assistance with coding from published sources was provided by S. Nordhoff and V. Rodrigues. We thank M. Mirazon-Lahr, E. Lindström, and G. Senft for discussion.

Phenotypic Diversity, Population Growth, and Information in Fluctuating Environments

Edo Kussell* and Stanislas Leibler

Organisms in fluctuating environments must constantly adapt their behavior to survive. In clonal populations, this may be achieved through sensing followed by response or through the generation of diversity by stochastic phenotype switching. Here we show that stochastic switching can be favored over sensing when the environment changes infrequently. The optimal switching rates then mimic the statistics of environmental changes. We derive a relation between the long-term growth rate of the organism and the information available about its fluctuating environment.

Organisms adapt readily to regularly varying environments, for instance, by adjusting to the daily light cycles by using internal circadian clocks. Real problems arise when environmental fluctuations are irregular. Organisms can adapt to sudden changes in chemical composition, local temperature, or illumination by sensing the changes and responding appropriately, for example, by switching phenotype or behavior. But there is a cost: each individual must maintain active sensory machinery.

Population diversity offers an alternate way to adapt to randomly fluctuating environments. Different subsets of the total population may be well-adapted to different types of environments. In genetically clonal populations, phenotypic diversity is generated by stochastic phenotype-switching mechanisms (*1–9*). Examples include flagellin phase variation in *Salmonella enterica* (*6*); microsatellite length variation (slipped-strand mispairing), controlling the expression of contingency genes in *Haemophilus influenzae* (*2, 4*); and swarming motility in *Bacillus subtilis* (*8*). The persistence mechanism in *Escherichia coli*, by which cells switch spontaneously and reversibly to a phenotype exhibiting slower growth and reduced killing by antibiotics (*9*), allows cells to survive prolonged exposure to antibiotics (*10*). Many other switching mechanisms are known in diverse bacteria (*2, 7*), fungi (*1–3*), and slime molds (*1*).

The idea that randomization of phenotype can be advantageous in fluctuating environments is well established in the ecology and population genetics literature (where it is known as bet-hedging). This idea has found applications in diverse contexts (*11*), and it was previously analyzed in several theoretical and computational studies (*12–18*).

We consider two extreme types of phenotype switching: responsive switching (*R*), occurring as a direct response to an outside cue detected by a sensing mechanism, and spontaneous stochastic switching (*S*), occurring without any direct sensing of the environment. Within a theoretical model, we address several questions. First, under which circumstances should each mechanism be used? For instance, if the detection of a sudden unfavorable environmental change, or the subsequent response, would be too slow, then it could be advantageous to have a subpopulation ready in an appropriate phenotype, before the environmental change.

Second, what determines parameters such as the switching rates? Random environmental

Laboratory of Living Matter and Center for Studies in Physics and Biology, The Rockefeller University, 1230 York Avenue, Box 34, New York, NY 10021–6399, USA.

*To whom correspondence should be addressed. E-mail: kussele@rockefeller.edu

fluctuations experienced by a population over long periods can select among different genetic mechanisms for generating diversity. Organisms whose stochastic switching rates are better adjusted to environmental fluctuations can outgrow organisms that use a different set of switching parameters.

Finally, how does information gathered by organisms about the fluctuating environment affect their survival? The two types of switching differ markedly in this regard. For responsive switching, information about environmental changes is conveyed to the organism directly through the sensing mechanism; whereas for stochastic switching, it is conveyed indirectly by natural selection.

We consider a simple model that encompasses both responsive and stochastic switching and describes a clonal population growing in an environment that fluctuates in time among a finite number (n) of different environment types (Fig. 1). The fluctuating environment is a continuous-time stochastic process, $\mathcal{E}(t)$, designating which environment occurs at time t; the average duration of environment i is τ_i (with the average over all environments equal to τ); the occurrence probability of environment i is p_i; and the probability that environment i follows j is b_{ij} ($b_{ii} = 0$).

Each individual organism is capable of exhibiting one of n different phenotypes. Phenotype i grows with rate $f_i^{(k)}$ in environment k (growth rates may be positive or negative). The phenotype with largest growth rate in environment k is phenotype k (its growth rate is $f_k^{(k)}$), and we refer to it as the fastest-growing phenotype and to all other phenotypes as slower phenotypes. Individuals may switch phenotype at any time, with parameters $H_{ij}^{(k)}$ giving the switching rate from phenotype j to phenotype i in environment k.

Taking the simplest model of growth, the n-dimensional population vector, $\mathbf{x}(t)$, whose ith coordinate is the number of individuals with phenotype i at time t, obeys the following equation

$$\frac{d}{dt}\,\mathbf{x}(t) = A_{\mathcal{E}(t)}\mathbf{x}(t)$$

The matrix $A_{\mathcal{E}(t)}$ may be one of n different matrices, depending on the environment, $\mathcal{E}(t)$. A_k can be written as a sum of a diagonal matrix, whose diagonal entries are the growth rates of each phenotype in environment k ($f_i^{(k)}$), and the matrix of switching rates, $H_{ij}^{(k)}$ (Fig. 1). The sum of all the entries of $\mathbf{x}(t)$ gives the total population size $N(t)$ (19).

The two types of phenotype switching correspond to different choices of switching rates. For stochastic switching, these rates are independent of the environment k; therefore, for all values k

$$H_{ij}^{(k)} = H_{ij} \text{ (stochastic switching)}$$

For responsive switching, the sensing mechanism allows switching rates to depend strongly on k. In the extreme case, all phenotypes switch with the same rate H_m to phenotype k in environment k, so

$$H_{kj}^{(k)} = H_m \text{ for all } j \neq k \text{ (responsive switching)}$$

$$H_{ij}^{(k)} = 0 \text{ for all } i \neq k \text{ and } j \neq i$$

The switching rate H_m is physiologically determined but ideally as large as possible, so that individuals spend as little time as possible in slower phenotypes.

To compare the two types of switching, we calculate the so-called Lyapunov exponent Λ (20), which is the asymptotic growth rate of total population size (21, 22) given by the large time limit of $(1/t) \log N(t)$. Λ is known to exist under relatively general conditions (20) and depends on both the organism (growth rates of its phenotypes and switching rates) and on the temporal sequence of the changing environment $\mathcal{E}(t)$. In general, it is difficult to compute analytically, but we now describe an approximation that allows such computation for our model.

We assume that environmental durations are long enough that the population has time to reach its equilibrium composition before the environment changes. In environment j, this means that $\mathbf{x}(t)$ will eventually point essentially in the direction of the top eigenvector of the matrix A_j. Upon a change of environment from j to i, there will be a delay time, T_{ij}^*, during which the population's composition changes
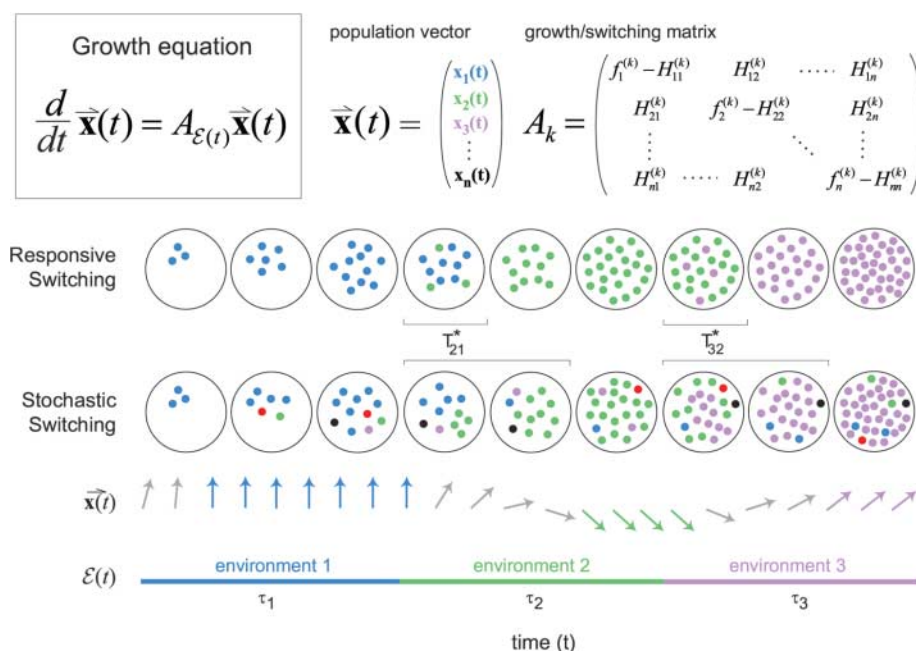

Fig. 1. A population is composed of individuals each capable of exhibiting one of n different phenotypes in n different environments. The growth rate of phenotype i in environment k is $f_i^{(k)}$; among all phenotypes in environment k, phenotype k grows the fastest. Individuals can switch phenotype at any time, responsively or stochastically. $H_{ij}^{(k)}$ is the switching rate from phenotype j to i in environment k, and $H_{jj}^{(k)} = \sum_{i \neq j} H_{ij}^{(k)}$. The boxed growth equation governs the dynamics of the n-dimensional population vector of phenotypes, $\mathbf{x}(t)$. The changing environment is a continuous-time stochastic process, $\mathcal{E}(t)$, taking integer values 1 to n designating the environment at time t. The form of matrices A_k is shown, determining the combined growth and switching rates of all phenotypes when $\mathcal{E}(t) = k$. $\mathcal{E}(t)$ is assumed to be constant on successive time intervals T_l with $l = 1, 2, \ldots$; $\varepsilon(l)$ is the environment occurring at the lth interval, and $L(t)$ denotes the number of intervals T_l elapsed by time t. Environment change probabilities are $b_{ij} \equiv P[\varepsilon(l) = i \mid \varepsilon(l-1) = j]$, so $\varepsilon(l)$ is a Markov chain (assumed ergodic) with n states and transition matrix b, with $b_{ii} = 0$. The equilibrium probability p_i of environment i satisfies $p_i = \sum_j p_j b_{ij}$; the average duration of environment i is τ_i and the average duration of environments is $\tau \equiv \sum_{i=1}^{n} p_i \tau_i$. Total population size is $N(t) \equiv \sum_{i=1}^{n} x_i(t)$. A schematic of the dynamics is shown in which individuals are colored to indicate phenotype, such that the fastest-growing phenotype in each environment matches the environment's color. When environment j changes to i, there is a delay time, T_{ij}^*, in which $\mathbf{x}(t)$ rotates (shown in gray) before the population attains its new composition. In responsive switching, individuals switch directly to the fastest-growing phenotype. In stochastic switching, subpopulations exist in different phenotypes; when the environment changes, the fastest-growing subpopulation brings about a change in population composition. Proportions of slower-growing phenotypes are exaggerated for the purpose of illustration; they may be as small as $\approx 10^{-6}$.

from its old structure (top eigenvector of A_j) to its new one (top eigenvector of A_i) (Fig. 1) (*23*). Thereafter, the population will grow at a rate given by the top eigenvalue of the matrix A_i, $\lambda_1(A_i)$. This simple picture allows for computation of the Lyapunov exponent in the limit of long durations (*24, 25*).

We find, in this limit, that Λ depends only on mean environmental durations τ_i and transition probabilities b_{ij} and is independent of other characteristics of environmental fluctuations (for example, the variance of environmental durations) (*24*). The biological implication is that a stochastic-switching organism is buffered against changes in the distribution of environmental variations, provided its environment does not fluctuate too quickly. We have verified this observation by simulation (fig. S1).

We now turn to the specific cases of responsive and stochastic switching. Because responsive switching requires a sensing apparatus in addition to the machinery of switching, we introduce the "cost of sensing" c to be the reduction of growth rates due to the presence of sensing machinery. We obtain the long-term growth rate for responsive switching

$$\tau\Lambda_R = \sum_{i=1}^{n} p_i\tau_i f_i^{(i)} -$$
$$c\tau - \sum_{i,j=1}^{n} p_j b_{ij}\, \log(1 + \Delta_{ji}^{R}/H_m)$$

[long-term growth] = [fastest growth] −
[sensing cost] − [delay-time cost]

and for stochastic switching (for small switching rates)

$$\tau\Lambda_S = \sum_{i=1}^{n} p_i\tau_i f_i^{(i)} -$$
$$\sum_{i=1}^{n} p_i\tau_i H_{ii} -$$
$$\sum_{i,j=1}^{n} p_j b_{ij}\, \log(1 + \Delta_{ij}^{S}/H_{ij}) + \dots$$

[long-term growth] = [fastest growth] −
[diversity cost] − [delay-time cost]

where $\Delta_{ij} \equiv f_j^{(j)} - f_i^{(j)}$, $\Delta_{ij}^{R} \equiv \Delta_{ij}$, and $1/\Delta_{ij}^{S} \equiv 1/\Delta_{ij} + 1/\Delta_{ji}$.

The general form of the expression for Λ is the same in both cases; only the origin of the second term is different. In stochastic switching, switching to slower phenotypes decreases Λ, incurring a "diversity cost"; in responsive switching, the "sensing cost" appears instead (*26*). The third term in both equations, the "delay-time cost," has a similar form, and is due to the time it takes the population structure to change after a change of environment.

We can now find the switching rates H_{ij} that maximize Λ for a stochastic switching organism, using the above expression for Λ_S

$$H_{ij}(\text{optimal}) = b_{ij}/\tau_j$$

The optimal switching rate from phenotype j to i is proportional to the probability that the environment changes from environment j to i and inversely proportional to the average duration of environment j. Optimal rates are thus precisely tuned to environmental statistics (*14, 27*).

The long-term growth rate for optimal switching is found to be

$$\tau\Lambda_S(\text{optimal}) = \sum_{i=1}^{n} p_i\tau_i f_i^{(i)} - 1 -$$
$$\sum_{i,j=1}^{n} p_j b_{ij}\, \log(\Delta_{ij}^{S}\tau_j) - I_{\text{env}}$$

where $I_{\text{env}} = -\sum_{i,j=1}^{n} p_j b_{ij}\, \log b_{ij}$. The term I_{env} is the entropy, or information content, of the fluctuating environment (*28*). It measures how unpredictable or surprising are the different environmental transitions appearing in the time sequence $\mathcal{E}(t)$. Even for optimal switching rates, the negative term $-I_{\text{env}}$ is present in Λ_S, because stochastic-switching organisms cannot perfectly anticipate the next environment (except when $I_{\text{env}} = 0$). In contrast to responsive switching, which senses a new environment, the stochastic-switching organism cannot overcome the entropy of its environment.

The appearance of I_{env} explicitly in the optimal long-term growth rate of a population points to possibly deeper connections between the fields of population biology and information theory (*29*). For example, consider a stochastic-switching organism with suboptimal rates H_{ij}'. These rates would be optimal in a varying environment with average durations $\tau_j' = 1/\sum_{k \neq j} H_{kj}'$ and transition probabilities $b_{ij}' = H_{ij}'/\sum_{k \neq j} H_{kj}'$. This organism has inaccurate information about its environment, as reflected in its long-term growth rate being lower than $\Lambda_S(\text{optimal})$ by $\frac{1}{\tau}S_{\text{K-L}} + \frac{1}{\tau}\sum_i p_i[\tau_i/\tau_i' - 1 - \log(\tau_i/\tau_i')]$. The relative entropy (Kullback-Leibler divergence), $S_{\text{K-L}} \equiv \sum_{i,j} p_j b_{ij}\log(b_{ij}/b_{ij}')$, is the penalty paid for poor information about environmental transitions (*30*).

We may also consider stochastic switching with memory, i.e., when individuals remember the last few phenotypic switches that occurred in their ancestral history. Such memory is advantageous when the fluctuating environment exhibits longer correlations, i.e., when environmental transition probabilities depend on the last m environments. As in the case of

sensing, there is a maximal cost for which memory is beneficial, which is related to the amount of information about the environment that such memory can reveal (*25*).

We now compare responsive and stochastic switching using long-term growth rates. If the cost of sensing, c, is small, responsive rather than stochastic switching will be favored. The maximal c for which sensing is advantageous is determined by the inequality $\Lambda_R > \Lambda_S(\text{optimal})$. We assume that fastest phenotypes have identical growth rates f, slower phenotypes have growth rates $\tilde{f}$ ($\Delta \equiv f - \tilde{f}$), and all environments have the same duration τ, and we find $c < \frac{1}{\tau}[1 + \log(\Delta\tau/2) - \log(1 + \Delta/H_m) + I_{\text{env}}]$.

The greater the uncertainty of the environment (I_{env}), and the faster the responsive organism responds (H_m), the higher the maximal cost c for which sensing is beneficial. In other words, an organism can afford to pay more for a sensor the more uncertain its environment. The longer environments remain constant, however, the less it pays to have a sensor. Stochastic switching is therefore favored when environments change infrequently (*31–34*).

The Lyapunov exponent can be defined even when total population size is bounded, and our main results apply in that case as well, provided that all slower phenotypes are represented in the population (*25*). We have not considered the evolutionary process that adjusts the switching rates. Given enough time, natural selection should change H_{ij} values toward the optimum, effectively extracting information about environmental statistics and translating it into switching rates (*35*).

We presented an analytical calculation of growth rates for structured populations in fluctuating environments and showed explicitly that information about environmental statistics is of central importance in population dynamics.

References and Notes
1. D. R. Soll, B. Kraft, *Dev. Genet.* **9**, 615 (1988).
2. E. R. Moxon, P. B. Rainey, M. A. Nowak, R. E. Lenski, *Curr. Biol.* **4**, 24 (1994).
3. J. Perez-Martin, J. A. Uria, A. D. Johnson, *EMBO J.* **18**, 2580 (1999).
4. C. D. Bayliss, D. Field, E. R. Moxon, *J. Clin. Invest.* **107**, 657 (2001).
5. S. A. Lachke, S. Joly, K. Daniels, D. R. Soll, *Microbiology* **148**, 2661 (2002).
6. H. R. Bonifield, K. T. Hughes, *J. Bacteriol.* **185**, 3567 (2003).
7. M. W. van der Woude, A. J. Baumler, *Clin. Microbiol. Rev.* **17**, 581 (2004).
8. D. B. Kearns, F. Chu, R. Rudner, R. Losick, *Mol. Microbiol.* **52**, 357 (2004).
9. N. Q. Balaban, J. Merrin, R. Chait, L. Kowalik, S. Leibler, *Science* **305**, 1622 (2004).
10. Whereas current experiments have not detected any direct sensing in the case of *E. coli* persistence, further experiments are needed to completely rule out the possibility of more subtle effects, such as the dependence of switching rates on the presence of antibiotics. Direct sensing of the environment might be involved in some of the known mechanisms that are thought to be purely stochastic, and in general, we may expect to find mixtures of stochastic switching and sensing mechanisms.

11. J. Seger, H. J. Brockmann, in *Oxford Surveys in Evolutionary Biology*, P. H. Harvey, L. Partridge, Eds. (Oxford Univ. Press, Oxford, 1987), vol. 4, pp. 182–211.

12. P. Haccou, Y. Iwasa, *Theor. Popul. Biol.* **47**, 212 (1995).

13. A. Sasaki, S. Ellner, *Evolution* **49**, 337 (1995).

14. M. Lachmann, E. Jablonka, *J. Theor. Biol.* **181**, 1 (1996).

15. A. Grafen, *Proc. R. Soc. London Ser. B* **266**, 799 (1999).

16. A. J. Robson, C. T. Bergstrom, J. K. Pritchard, *J. Theor. Biol.* **197**, 541 (1999).

17. M. Thattai, A. van Oudenaarden, *Genetics* **167**, 523 (2004).

18. D. M. Wolf, V. V. Vazirani, A. P. Arkin, *J. Theor. Biol.* **234**, 227 (2005).

19. Here populations are taken to be large enough that deterministic dynamics of $x(t)$ can be assumed. Stochastic dynamics of small populations are discussed briefly in (*25*).

20. A. Crisanti, G. Paladin, A. Vulpiani, *Products of Random Matrices* (Springer-Verlag, Berlin, 1993).

21. R. C. Lewontin, D. Cohen, *Proc. Natl. Acad. Sci. U.S.A.* **62**, 1056 (1969).

22. J. E. Cohen, *Stochast. Process. Appl.* **9**, 245 (1979).

23. For small switching rates, the delay times can very roughly be thought of as the time for phenotype i to sweep through the population and replace phenotype j as the dominating type. For large switching rates, this is not the case, because the population can be a mixture of multiple phenotypes having comparable numbers of individuals.

24. Following the notation of Fig. 1 and its caption, growth in each environment will, after some delay time T_l^*, be driven solely by the fastest phenotype at a rate $\lambda_1(A_{\varepsilon(l)})$. Provided the environmental duration T_l is long, population growth due to slower phenotypes during the delay time may be ignored with negligible error, yielding $\Lambda = \lim_{t \to \infty} \frac{1}{t} \sum_{l=1}^{L(t)} \lambda_1(A_{\varepsilon(l)})(T_l - T_l^*)$. The

delay times T_l^* depend only on environments $\varepsilon(l)$ and $\varepsilon(l-1)$, so $T_l^* = T_{\varepsilon(l)\varepsilon(l-1)}^*$. Because $t \to \tau L(t)$, we find $\Lambda = \frac{1}{\tau} \sum_{i,j=1}^{n} p_j b_{ij} \lambda_1(A_i)(\tau_i - T_{ij}^*)$. Calculation of $\lambda_1(A_i)$ and T_{ij}^* then gives Λ.

25. Details of our calculation, the approximations used, and numerical computations are available as supporting material on *Science* Online.

26. In deriving these expressions, we have assumed that environmental durations are sufficiently long and that switching rates are strictly positive. For Λ_s, we also assumed that switching rates H_{ij} are small (*25*).

27. In certain cases, not switching at all, i.e., remaining in a single phenotype at all times, is better than switching phenotype. This can be the case, for example, when environments change very rapidly. When environments last sufficiently long, the solution given in the text is the optimum.

28. T. M. Cover, J. A. Thomas, *Elements of Information Theory* (Wiley, New York, 1991).

29. The notion of entropy was used previously in population biology as an empirical measure of population diversity, and in the context of age-structured populations as a tool for solving the dynamic equation (*36, 37*). Demetrius observed that several properties of the stable age distribution are determined by an entropy defined on population genealogies.

30. A similar result is seen in certain types of gambling (*28*), e.g., horse races and stock markets, where the gambler must decide how much money to bet on each possible outcome.

31. R. B. Bourret, A. M. Stock, *J. Biol. Chem.* **277**, 9625 (2002).

32. J. M. G. Vilar, C. C. Guet, S. Leibler, *J. Cell Biol.* **161**, 471 (2003).

33. E. Kussell, R. Kishony, N. Q. Balaban, S. Leibler, *Genetics* **169**, 1807 (2005).

34. Many sensing systems exist in microorganisms such as those involved in chemotaxis (*31*), the *lac* operon (*32*), and onset of stationary phase. In these examples and others, the environmental cue that is sensed occurs often, on the order of tens of generations or less, and our model predicts that sensing is advantageous. For bacterial persistence, experiments suggest that direct sensing might be absent (*9*). The appearance of antibiotic may be a relatively rare event in the natural habitats of *E. coli* (although direct evidence is scarce), in which case our model favors stochastic switching (*33*).

35. By comparing the time to fixation for a mutation that abolishes phenotype switching with environmental durations, we find that natural selection can maintain switching mechanisms. The mechanism may be lost if a very long environmental duration ($\gg \tau_i$) occurs (*25*).

36. L. Demetrius, *J. Stat. Phys.* **30**, 709 (1983).

37. P. Michel, S. Mischler, B. Perthame, *C. R. Acad. Sci. Paris* **338**, 697 (2004).

38. We would like to thank G. Ben Arous, N. Balaban, R. Kishony, C. Bargmann, J. McKinney, L. Keller, R. Losick, A. Murray, A. Grosberg, J. Merrin, and O. Ahmed for discussions. We particularly thank G. Ben Arous and C. Bargmann for their encouragements and many detailed comments about the manuscript. E.K. thanks the Alfred P. Sloan Foundation for financial support.