

# Genetic drift at expanding frontiers promotes gene segregation

Oskar Hallatschek<sup>\*†</sup>, Pascal Hersen<sup>†‡</sup>, Sharad Ramanathan<sup>†§</sup>, and David R. Nelson<sup>\*†¶</sup>

<sup>\*</sup>Department of Physics and <sup>†</sup>FAS Center for Systems Biology, Harvard University, Cambridge, MA 02138; <sup>‡</sup>Centre National de la Recherche Scientifique, Laboratoire Matière et Systèmes Complexes, Université Paris VII, 75013 Paris, France; and <sup>§</sup>Bell Laboratories, Alcatel-Lucent, Murray Hill, NJ 07974

Contributed by David R. Nelson, October 25, 2007 (sent for review September 7, 2007)

**Competition between random genetic drift and natural selection play a central role in evolution: Whereas nonbeneficial mutations often prevail in small populations by chance, mutations that sweep through large populations typically confer a selective advantage. Here, however, we observe chance effects during range expansions that dramatically alter the gene pool even in large microbial populations. Initially well mixed populations of two fluorescently labeled strains of *Escherichia coli* develop well defined, sector-like regions with fractal boundaries in expanding colonies. The formation of these regions is driven by random fluctuations that originate in a thin band of pioneers at the expanding frontier. A comparison of bacterial and yeast colonies (*Saccharomyces cerevisiae*) suggests that this large-scale genetic sectoring is a generic phenomenon that may provide a detectable footprint of past range expansions.**

genetic distance | founder effect | microorganisms | range expansion

A principal tenet of modern evolutionary biology is that Darwinian selection and random genetic drift compete in driving evolutionary change (1). It is widely accepted that genetic drift can have significant effects on small populations (2) that may even lead to speciation (3). In large populations, however, random sampling effects are generally considered weak compared with selection (4–6) (law of large numbers). A major departure from this paradigmatic behavior occurs when large populations undergo range expansions. The descendants of individuals first settling in a new territory are most likely to dominate the gene pool as the expansion progresses (7–9). Random sampling effects among these pioneers results in genetic drift that can have profound consequences on the diversity of the expanding population. Indeed, spatially varying levels of genetic diversity and colonization patterns appear to be correlated in many species (10–13). For example, the often observed south–north gradient in neutral genetic diversity [“southern richness to northern purity” (14)] on the northern hemisphere is thought to reflect past range expansions induced by glacial cycles (12). Although these trends indicate that genetic drift during range expansions has shaped the gene pool of many species, the underlying spatial mechanism remains obscure: Diversity gradients are often difficult to interpret and potentially interfere with the signal of spreading beneficial mutations (15, 16). In fact, a major challenge of present-day population genetics is to decide whether natural selection or a past demographic process is responsible for the prevalence of common mutations (17).

Here, we use simple microbial systems to study the nature of random genetic drift in range expansions of large populations. We observe chance effects that segregate the gene pool into well defined, sector-like regions of reduced genetic diversity. The genetic segregation on the population level is the consequence of number fluctuations on a much smaller scale, within a thin region of reproducing pioneers at the expanding frontier. We expect these patterns to be a general signature of continuous range expansions in populations exhibiting moderate rates of turnover and migration. Our results have significant implications

for the identification and demographic interpretation of neutral mutations that swept through a population merely by chance.

## Results and Discussion

We used two strains of fluorescently labeled bacteria (*Escherichia coli*) to track both the neutral gene dynamics and the population growth during a range expansion. The strains were genetically identical except for expressing either cyan fluorescent protein (CFP) or yellow fluorescent protein (YFP), which differ only by a single point mutation, from the same constitutively active promoter (see *Materials and Methods*). We mixed both strains to obtain cultures with different proportions and placed a droplet containing cells at the center of an agar plate containing rich growth medium (LB). Population growth and mutant distribution were monitored for 4 days by means of a stereo microscope. Fig. 1 summarizes the colonial development of a founder population of CFP- to YFP-tagged strains at mixing ratios of 50:50 and 95:5. The time series of fluorescence images in Fig. 1A shows that, in our experiments with immotile *E. coli* on hard agar, there is no noticeable temporal change in the mutant distribution behind the population front. The mutant distribution evolves only at the leading edge, much like a (dichromatic) carpet that is knitted from the inside to the outside. The final state of this “genetic” carpet (Fig. 1B and C) thus can be taken as a frozen record of the colonization process.

The observed mutant distributions can be partitioned into areas of two very different patterns. The central region of Fig. 1B, which we refer to as the homeland, exhibits a dense speckle pattern reminiscent of the homogeneous and well mixed initial population. This homeland is bounded by a ring of higher speckle density, which was formed at the edge of the initial droplet (innermost black line in Fig. 1B) during evaporation of its water content [most likely due to flows generated by the evaporation process (18), see [supporting information \(SI\) Fig. 5](#)]. From this ring toward the boundary of the mature colony, the population segregates into single-colored domains. A large number of differently labeled flares radiate and gradually coalesce into a few large sectors that reach the edge of the mature colony. Fig. 1C, which corresponds to the unbalanced mixture 95:5 of YFP to CFP, also shows a subdivision into an inner and outer region. The main difference is that the spots and flares are so dilute that their evolution is effectively “noninteracting” and can be individually identified. In particular, only two spots of the minority population evolve into sectors that actually reach the edge of the mature colony.

The formation of morphological sectors, often visible to the naked eye, is a well known phenomenon in microbial colonies

Author contributions: O.H., S.R., and D.R.N. designed research; O.H., P.H., S.R., and D.R.N. performed research; O.H. and D.R.N. analyzed data; and O.H. wrote the paper.

The authors declare no conflict of interest.

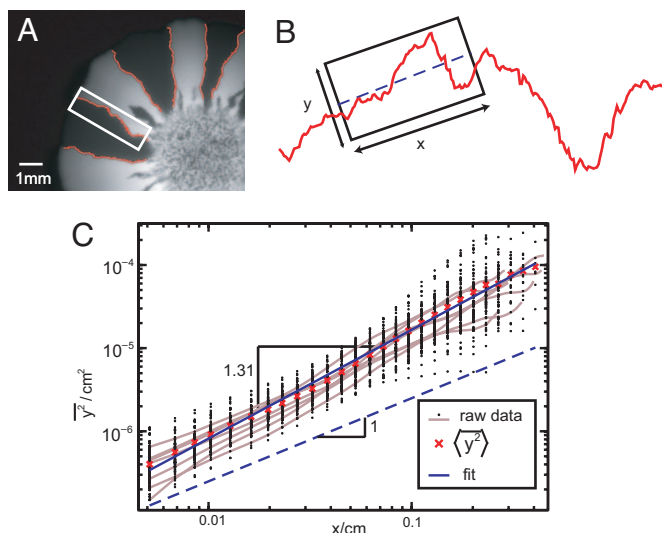
Freely available online through the PNAS open access option.

<sup>¶</sup>To whom correspondence should be addressed. E-mail: nelson@cmt.harvard.edu.

This article contains supporting information online at [www.pnas.org/cgi/content/full/0710150104/DC1](http://www.pnas.org/cgi/content/full/0710150104/DC1).

© 2007 by The National Academy of Sciences of the USA

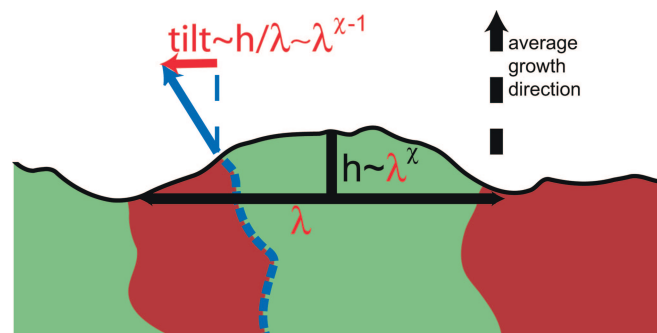




**Fig. 2.** Random meandering of domain boundaries is superdiffusive. (A) From fluorescent images of *E. coli* colonies, domain boundaries were extracted (red lines) by using an edge-detection algorithm. (B) To test whether these domain boundaries resemble time traces of diffusive random walks, we quantified their contour fluctuations as a function of contour length. To this end, we chose a sliding window of size  $x$ , parameterized the contour by the transverse displacements  $y$  from the linear best fit, and determined the mean square displacement  $\overline{y^2}$  (see *Materials and Methods*). (C) The mean square transverse displacement is shown as a function of the size  $x$  of the sliding window. Ninety-two domain boundaries sampled from 12 different *E. coli* colonies were analyzed (black points). Even though the data for individual domain boundaries (gray lines) are quite noisy, the ensemble average (red crosses) very closely follows a straight line in this double logarithmic plot. The exponent  $2\zeta = 1.31 \pm 0.1$  of the fitted power law (blue line) indicates that boundaries carry out superdiffusive random walks around their average growth direction, i.e., they wander more vigorously than conventional random walks, which have a wandering exponent given by  $2\zeta = 1$ .

leading tip of the domain wall drifts transverse to the growth direction, as illustrated in Fig. 3. This drift fluctuates in time, just as the shape of the colonial front does, and has zero mean. By providing a fluctuating bias, a time-dependent interface roughness is thus expected to amplify the random wandering of the domain boundaries. The following scaling argument suggests that this effect can even change the character of the random walk: Suppose the interface kinetics of the colony is characterized (25) by a roughness exponent (25)  $\chi$  and a dynamical exponent  $\zeta$ ; i.e., an interface segment of linear dimension  $L$  exhibits displacements of the order  $L^\zeta$  from a straight line and relaxes on a time scale  $L^\zeta$ . We then argue that, during the time  $t$ , a transverse drift  $\propto L(t)^{\chi-1} \sim t^{(\chi-1)/z}$  acts on the leading tip of a domain wall that is proportional to the tilting angle of a typical undulation on the scale  $L$ . This drift entails transverse displacements of the random walk scaling like  $t^{1+(\chi-1)/z}$ , suggesting a wandering exponent of  $\zeta = 1 + (\chi - 1)/z$ . Our deterministic analysis is valid only in the case of superdiffusion, such that any intrinsic regular diffusion is negligible on sufficiently large times. A simple bacterial growth model, the Eden model<sup>11</sup> (26), is known to be characterized by  $\chi = 1/2$  and  $z = 3/2$  in one dimension, leading to a wandering exponent of  $\zeta = 2/3$  [verified by simulations (27)], which is consistent with the exponent  $\zeta = 0.65 \pm 0.05$  found in our experiments (Fig. 2C).

Because the endpoints of domain walls carry out random walks (with the radial axis representing time) and annihilate when they collide, their gradual decimation may be described by a one-

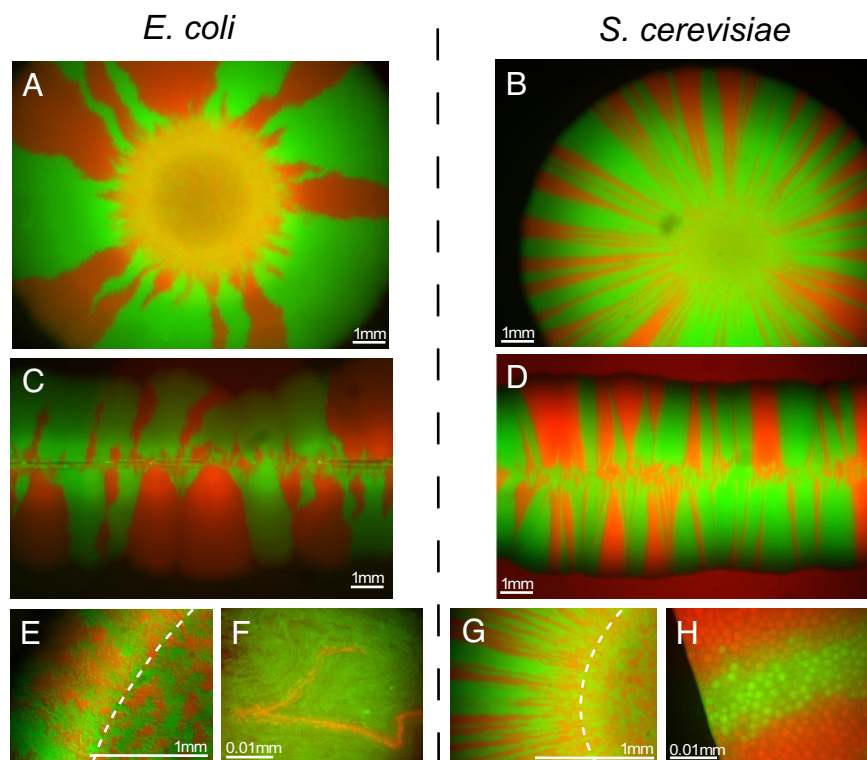


**Fig. 3.** The roughness of the colonization front (black line) influences the wandering of locally perpendicular domain boundaries (red lines). In this sketch, the middle domain boundary (dashed blue) tends to follow the blue arrow indicating the local growth direction of the colony. Because of the stochastic surface growth, the local growth direction deviates from the average growth direction (black dashed arrow). Consequently, the domain boundary is subject to a drift (red arrow) transverse to the average growth direction, which is proportional to the local tilt of the interface. This tilt depends on the roughness of the edge of the colony, which we assume to be characterized by a roughness exponent  $\chi$ : On a length scale  $\lambda$ , the edge of the colony exhibits stochastic shape undulations of amplitude  $h \sim \lambda^\chi$ . The corresponding tilt,  $h/\lambda \sim \lambda^{\chi-1}$ , biases the path of the considered domain boundary until the surface irregularity causing it has relaxed. Because of the roughness of the colonial edge, the domain boundaries thus experience a fluctuating bias, which may significantly change their wandering statistics, as explained in the text.

dimensional model of annihilating random walkers (28) (with periodic boundary conditions). These random walkers, which come in pairs, annihilate completely for large enough times, if their available space is finite. This scenario applies to a colony that advances approximately as a straight line, as can be induced experimentally through a line inoculation (compare Fig. 4C and D). For describing the case of circular microbial colony, such as Fig. 1, this analysis has to be modified to account for the fact that the circumference grows linearly in time. Because of this inflation, two domain walls tend to drift apart at a rate proportional to the velocity of the expanding population front. Incorporating this feature into an annihilating walker model with inflation reveals a cross-over time beyond which the increase in the size of the circumference is faster than the random walk of the domain walls. As a consequence, domain walls stop colliding, and their mutual annihilation ends. Thus, domain walls have a nonzero survival probability in circular colonies. A simple scaling argument balancing the opposing effects of inflation and domain-wall wandering indicates that the expected genetic diversity, as reflected by the sectoring pattern, is reminiscent of the shape of the precolonial habitat rather than the initial number of founder cells. In particular, we expect the relationship  $N_s \sim r_0^{1-\zeta}$  between the average number  $N_s$  of surviving sectors and the radius  $r_0$  of the homeland (see *SI Text*). The enhanced wandering ( $\zeta \approx 0.65 > 1/2$ ) of domain walls amplifies the effects of genetic drift and leads to a more rapid decrease in genetic diversity than expected for a purely diffusive wandering.

Our neutral model for the coarsening dynamics in expanding *E. coli* colonies suggests that the sectoring phenomenon is not unique to the considered microbial species but caused by the nature of chance effects during colonization. Indeed, our experiments on growing colonies of haploid *Saccharomyces cerevisiae* give qualitatively similar results (Fig. 4). Again, we observed an initial coarsening phase, in which domain boundaries annihilate, followed by a stationary phase, in which the sectoring pattern is stable. Compared with *E. coli*, however, yeast colonies have a much larger number of surviving sectors and much straighter domain boundaries. Both observations are consistent with our model of annihilating domain boundaries, because the frequency

<sup>11</sup>Eden M (1961) Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, June 20–30, 1960, Berkeley, CA, 4:233.



**Fig. 4.** Segregation patterns of two different microbial species are qualitatively similar but different in detail. (A and B) Both bacteria (*E. coli*) and yeast (*S. cerevisiae*) colonies exhibit spatial gene segregation when they grow colonies on agar plates. However, the number of (surviving) sectors is much larger in yeast colonies. For the linear inoculations (C and D), the Petri dish was gently touched with a sterile razor blade that was previously wetted by a liquid culture of a binary mixture of mutants. (E and G) Continuous patches of boundary regions and homeland (bounded by dashed line) at a magnification of  $\times 51$  for *E. coli* (E) and yeast (G). (F and H) Images at single-cell resolution ( $\times 100$ ). (F) Tip of a sector dies out (*E. coli*). (H) Section boundary at the frontier (yeast).

of annihilation events is expected to decrease for straighter domain boundaries. The disparate domain wall wandering could be due to differences in the mode of replication, the shape of the cells, or the thickness of the colonies. However, even though the local genetic patterns in colonies of different microbial species might be quite different, the genetic segregation on larger scales seems to be a more general consequence of a microbial range expansion.

### Conclusions

The observations presented here strongly suggest that enhanced genetic drift at expanding frontiers can dramatically alter the genetic makeup of large microbial populations. Neutral mutants quickly segregate into monoclonal domains that further coarsen as the colony grows. This coarsening process is driven by the rugged, fractal path of the domain boundaries. The final sectoring of circular colonies is controlled by a balance between the deterministic inflation of the colonial perimeter, dominant at large times, and the stochastic meandering of domain boundaries causing them to annihilate on short times.

More generally, our experiments demonstrate that mutations at expanding frontiers can sweep through a population without any selective advantage. This observation supports theoretical arguments and genetic evidence (15, 29) that common alleles in a population might not necessarily reflect positive selection but, instead, recent range expansions, which have probably occurred in many species (11, 12). Our characterization of the resulting sector-like regions of reduced genetic diversity may serve as a basis for discriminating these surfing mutations from sweeping beneficial mutations. The possibility of superdiffusive wandering (identified here for *E. coli*) would have to be considered in such a discrimination.

In light of inferring past migrations from spatially resolved genetic data, our results suggest that, apart from a much anticipated general reduction of genetic diversity by a range expansion, a fragmentation of the colonized regions into sectors has to be considered, which is stabilized during expansion by an inflationary effect. Although demonstrated here only for non-motile microbial species, we believe that the underlying segregation mechanism is quite generic for populations growing continuously and isotropically in two dimensions. As long as these genetic patterns are not blurred by subsequent dispersal and population turnover, they provide a record of the colonization dynamics and information about the precolonial habitat.

### Materials and Methods

**Bacterial Strains and Plasmids.** Our strains of immotile *E. coli* are based on a DH5 $\alpha$  background and contain plasmids expressing enhanced (e)CFP (A206K) and Venus YFP (A206K), respectively. These strains are identical except for the fluorescent protein structures: Genes for those proteins are cloned between the SacI and XbaI sites of the vector pTrc99A, which is Ampicillin resistant and isopropyl  $\beta$ -D-thiogalactoside (IPTG) inducible. The A206K mutation eliminates protein–protein aggregation. DH5 $\alpha$  showed sufficient background level expression even in the absence of IPTG. This expression allowed us to visualize the cells in the absence of any IPTG in the medium. Measurements of the exponential growth rate of both ancestral strains, and of isolates derived from sectors, indicate that the fitness differences are  $<5\%$  (see SI Fig. 6).

**Bacterial Growth Conditions.** The two strains of *E. coli* were grown overnight in separate liquid cultures at 37°C in rich growth medium (LB) supplemented with Ampicillin. After tenfold

dilution in fresh medium, we mixed them to obtain cultures with different proportions of the two kinds of cells. Droplets of different sizes were then pipetted on LB-agar plates (1.5% Agar) with Ampicillin, and grown at 37°C for up to 84 h. Population growth and mutant distribution were monitored at regular time intervals by means of a stereo microscope. No noticeable dependence of the final mutant distribution on the growth phase of the cells at inoculation time could be detected. The cell density used for the inoculation influenced the patchiness of the homeland, but had no apparent effect on the sectoring pattern.

**Yeast Strains and Plasmids.** The haploid strains (mating type) of *S. cerevisiae*, obtained from the E. O'Shea laboratory (Harvard University), are derived from a K699 background and have eCFP or Venus (YFP) driven by the ADH1 promoter integrated into the leu2 locus.

**Yeast Growth Conditions.** The two strains were grown overnight in separate liquid cultures at 30°C in yeast extract/peptone/dextrose (YPD) medium. After tenfold dilution in fresh medium, we mixed them to obtain cultures with different proportions of the two kinds of cells. Droplets of different sizes (1–8  $\mu$ m) were then pipetted on YPD-agar plates, and grown at 30°C for up to 5 days before microscopy.

**Microscopy and Image Processing.** The cells were observed by using a Zeiss fluorescent binocular microscope (Zeiss) and with a Zeiss Axiovert 200M (Zeiss), both with a Hamamatsu ORCA-AG camera (Hamamatsu). Emission from YFP was visualized at 535 nm (30-nm bandwidth) upon excitation at 500 nm (20-nm bandwidth), and emission of CFP was visualized at 470 nm (30-nm bandwidth) upon excitation at 430 nm (25-nm bandwidth). The image overlays in Figs. 1 *B* and *C* and 4 were created by using Adobe Photoshop (Adobe Systems). The fluorescence channels complement each other without any notice-

able gaps (i.e., no loss of plasmids). For Fig. 2*C*, the contours of 92 domain boundaries (of typical length 0.5 cm) from 18 repeated experiments were extracted on a resolution of 5  $\mu$ m by using the edge-detection algorithm of the software ImageJ (<http://rsb.info.nih.gov/ij/>). A C++ code was used to quantify the wandering of a given section (length  $L$ ) of a domain boundary in the following way. First, a longitudinal axis was defined by fitting a straight line to the considered contour (best linear fit). Second, the contour was parameterized in terms of its displacement transverse to the longitudinal axis  $y(x)$  as a function of the longitudinal coordinate  $x \in (0, L)$  (see Fig. 2*B*). Third, we computed the mean square displacement  $\langle y^2(L) \rangle = \int_0^L dx y^2(x)/L$ . To improve statistics, we averaged this quantity over the possible ways to define a section of length  $L$  within a domain boundary of larger length (running average). There are various other ways to quantify the wandering, which should all give the same asymptotic scaling. Because, however, the range of our data are limited to two orders of magnitude (see Fig. 2*C*), the fitted exponent somewhat depends on such a definition. We estimated the error in the exponent  $\zeta \approx 0.65 \pm 0.05$  by probing its dependence on the quantification method. This source of error is, by far, larger than the standard deviation of the ensemble average in Fig. 2*C* (red crosses) from the fitted line (blue).

We thank J. Shapiro and T. Shimizu for helpful discussions; T. Shimizu (Harvard University) together with H. Berg (Harvard University) and V. Sourjik (Center for Molecular Biology Heidelberg, ZMBH, Heidelberg, Germany) for the *E. coli* strains and plasmids; the O'Shea laboratory for *S. cerevisiae* strains; and M. Bathe, K. Foster, M. Nowak, B. Stern, J. Wakeley, and J. Xavier for providing comments on the manuscript. This work was supported by German Research Foundation Grant Ha 5163/1 (to O.H.), National Science Foundation through Grant DMR-0654191, the Harvard Materials Research Science and Engineering Center through Grant DMR-D2138D5 (to D.R.N.), a Human Frontiers Grant (to S.R.), a Keck Futures Initiative Grant (to S.R.), and a National Institute of General Medical Sciences Center grant (to P.H. and S.R.).

1. Nei M (2005) *Mol Biol Evol* 22:2318–2342.
2. Nei M, Maruyama T, Chakraborty R (1975) *Evolution (Lawrence, Kans)* 29:1–10.
3. Mayr E (1963) *Animal Species and Evolution* (Belknap, Cambridge, MA).
4. Elena SF, Lenski RE (2003) *Nat Rev Genet* 4:457–469.
5. Hegreness M, Shoresh N, Hartl D, Kishony R (2006) *Science* 311:1615–1617.
6. Hartl DL, Clark AG (1997) *Principles of Population Genetics* (Sinauer, Sunderland, MA).
7. Edmonds CA, Lillie AS, Cavalli-Sforza LL (2004) *Proc Natl Acad Sci USA* 101:975–979.
8. Klopstein S, Currat M, Excoffier L (2006) *Mol Biol Evol* 23:482–490.
9. Hallatschek O, Nelson DR (2007) *Theor Popul Biol*, doi:10.1016/j.tpb.2007.08.008.
10. Cavalli-Sforza LL, Menozzi P, Piazza A (1993) *Science* 259:639–646.
11. Cavalli-Sforza LL, Feldman MW (2003) *Nat Genet* 33:266–275.
12. Hewitt G (2000) *Nature* 405:907–913.
13. Ramachandran S, Deshpande O, Roseman CC, Rosenberg NA, Feldman MW, Cavalli-Sforza LL (2005) *Proc Natl Acad Sci USA* 102:15942–15947.
14. Hewitt GM (1996) *Biol J Linnean Soc* 58:247–276.
15. Currat M, Excoffier L, Maddison W, Otto SP, Ray N, Whitlock MC, Yeaman S (2006) *Science* 313:172a.
16. Schroeder KB, Schurr TG, Long JC, Rosenberg NA, Crawford MH, Tarskaia LA, Osipova LP, Zhadanov SI, Smith DG (2007) *Biol Lett* 3:218–223.
17. Pennisi E (2007) *Science* 316:1690–1692.
18. Deegan RD, Bakajin O, Dupont TF, Huber G, Nagel SR, Witten TA (1997) *Nature* 389:827–829.
19. Shapiro JA (1995) *BioEssays* 17:597–607.
20. Ben-Jacob E (2003) *Philos Trans R Soc London Ser A* 361:1283–1312.
21. Koshland D, Kent JC, Hartwell LH (1985) *Cell* 40:393–403.
22. Murray AW, Szostak JW (1983) *Cell* 34:961–970.
23. McMurray MA, Gottschling DE (2003) *Science* 301:1908–1911.
24. Berg HC (1993) *Random Walks in Biology* (Princeton Univ Press, Princeton).
25. Halpin-Healy T, Zhang Y (1995) *Phys Rep* 254:215–414.
26. Kardar M, Parisi G, Zhang YC (1986) *Phys Rev Lett* 56:889–892.
27. Saito Y, Mueller-Krumbhaar H (1995) *Phys Rev Lett* 74:4325–4328.
28. Toussaint D, Wilczek F (1983) *J Chem Phys* 78:2642–2647.
29. Biek R, Henderson JC, Waller LA, Rupprecht CE, Real LA (2007) *Proc Natl Acad Sci USA* 104:7993–7998.