

Reproductive pair correlations and the clustering of organisms

W. R. Young*, A. J. Roberts† & G. Stuhne*

* Scripps Institution of Oceanography, University of California at San Diego, La Jolla, California 92093-0230, USA

† Department of Maths & Computing, University of Southern Queensland, Toowoomba, Queensland 4352, Australia

Clustering of organisms can be a consequence of social behaviour, or of the response of individuals to chemical and physical cues¹. Environmental variability can also cause clustering: for example, marine turbulence transports plankton^{2–8} and produces chlorophyll concentration patterns in the upper ocean^{9–11}. Even in a homogeneous environment, nonlinear interactions between species^{12–14} can result in spontaneous pattern formation. Here we show that a population of independent, random-walking organisms ('brownian bugs'), reproducing by binary division and dying at constant rates, spontaneously aggregates. Using an individual-based model, we show that clusters form out of spatially homogeneous initial conditions without environmental variability, predator–prey interactions, kinesis or taxis. The clustering mechanism is reproductively driven—birth must always be adjacent to a living organism. This clustering can overwhelm diffusion and create non-poissonian correlations between pairs (parent and offspring) or organisms, leading to the emergence of patterns.

Figure 1 shows a simulation of the brownian bug model using an individual-based Monte Carlo approach (described below in Methods). Because the motion of a brownian bug is unaffected by a reproductive event, or by the proximity of other bugs, the spatial aspects of the model are only weakly coupled to the primitive biological ingredients of birth and death. Perhaps, then, it is surprising that continuum approximations do not represent the patches that form in the individual-based model shown in Fig. 1. The continuum approximation^{1,12} describes the evolution of large populations using advection–diffusion–reaction (ADR) equations for the concentration: $C(\mathbf{x}, t)dA$ is the expected number of organisms in the sample area dA surrounding the point \mathbf{x} at time t . The ADR approximation of the brownian bug process is:

$$C_t = D\nabla^2 C + (\lambda - \mu)C \quad (1)$$

where λ is the birth rate, μ is the death rate and D is the diffusivity. However, in Fig. 1 death and birth are equally probable so $\lambda = \mu$, and equation (1) collapses to the diffusion equation. Then, as the initial concentration C_0 is uniform, the solution of equation (1) is $C(\mathbf{x}, t) = C_0$; this is not a good characterization of Fig. 1. This failure shows that equation (1), and other ADR approximations^{2–8,13–15}, do not capture the fluctuations exhibited by the individual-based model in Fig. 1.

In reality there is a fundamental difference between the locations of birth and death: deaths occur anywhere, but birth always occurs adjacent to a living organism. This asymmetry is not represented in the continuum approximation of equation (1), in which the birth rate and the death rate occur only in the combination $\lambda - \mu$. These reproductive pair correlations are a uniquely biological complication, with no analogue in the physical and chemical problems that served as the initial inspiration¹⁵ for biological ADR modelling. Careful derivations of the ADR approximation¹² make the assumption that on the length scale of the sample area dA individuals are distributed without correlations. This assumption enables us to use Poisson statistics to calculate averages. The brownian bug process realized in Fig. 1 shows that this Poisson hypothesis can fail in simple circumstances.

Of course, uniform concentration is the correct (but useless) answer if $C(\mathbf{x}, t)$ is defined by an ensemble average over many realizations: the concentration fluctuations in Fig. 1 disappear after an ensemble average because the brownian bug process is spatially homogeneous. However, as discussed above, the concentration is usually defined via a spatial average over dA . The optimal choice of sampling scale is important^{16–18} because in most circumstances only spatial averages are accessible to observation, whereas the ensemble average is most useful for theoretical purposes. If the spatial average of a single realization is approximated by the ensemble average of many realizations (for example, either molecularly or turbulently diffusing chemicals) we say that the system is self-averaging. The

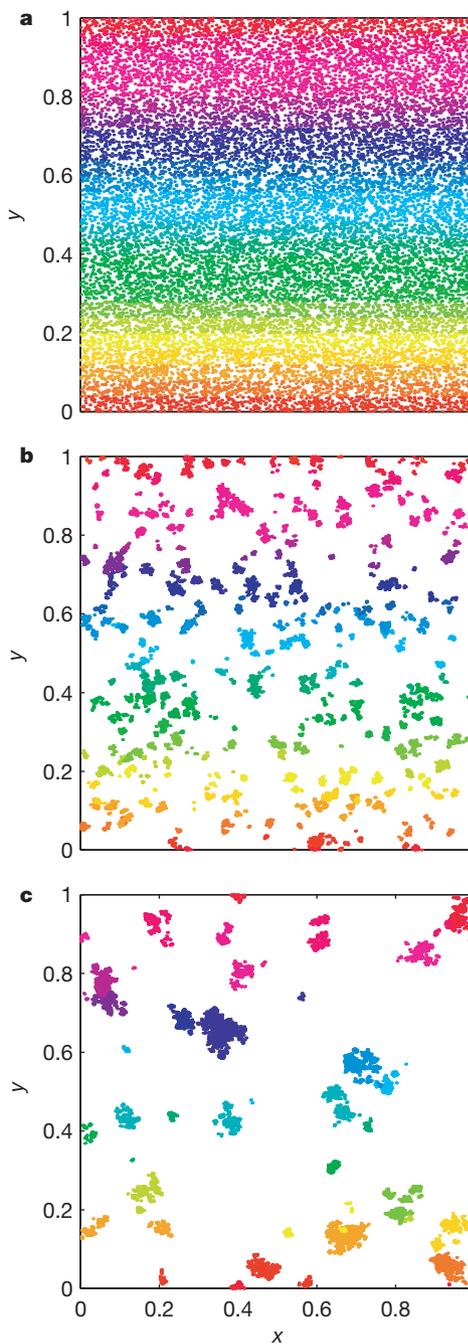


Figure 1 Distributions of brownian bugs at different times in a simulation with $\Delta = 10^{-3}$ and $N_0 = 20,000$. Each point is the position of a brownian bug in the (x, y) –plane. See Methods for the details of the simulation. **a**, Initial condition (a Poisson spatial distribution), with brownian bug colour coded by y coordinate. **b**, Descendants of the original brownian bugs at $t = 100\tau$. **c**, As **b**, at $t = 1,000$.

disagreement between the solution $C(\mathbf{x}, t) = C_0$ of equation (1) and the realization in Fig. 1 indicates a failure self-averaging.

For oceanographic systems it is interesting to study the effects of large-scale random advection on the patches in Fig. 1. This is the simplest model of a planktonic species reproducing in a turbulent fluid. The assumption that the flow field is large scale means that the characteristic length of the velocity is much greater than the diameter of the emerging patches in Fig. 1b. In other words, we are in the ‘Batchelor regime’ in which the velocity field varies linearly with the separation $s(t)$ of a recently divided parent and offspring^{19–22}. In this case there is a timescale (related to the root mean square, r.m.s., strain rate of the velocity) but there is no length scale, other than $s(t)$. Consequently $s(t) \propto \exp(\gamma t)$, where γ^{-1} is the timescale mentioned above.

Because exponential pair separation is much faster than diffusive separation, we might expect that the patches formed by reproductive pair correlations would disappear. However, the simulation in Fig. 2, which uses a computationally tractable model of exponential pair separation²², shows that the patches persist as elongated filaments (Fig. 2b). This elongation is suggested by visualizations of nonreproducing contaminants in this same incompressible random velocity field²² (Fig. 2a). Without reproduction an initially uniform density of brownian bugs remains uniform (Fig. 2a): stirring alone does not create patches. But with reproduction, filamentary patches appear spontaneously (Fig. 2b). Thus, there is an essential difference between reproductive and nonreproductive contaminants.

To proceed with theory beyond equation (1) we notice that the constant ensemble-averaged concentration C is the first member of a

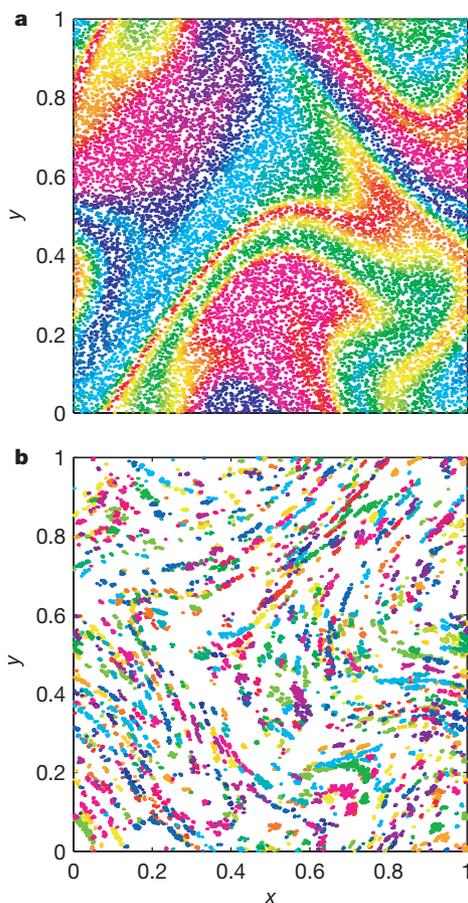


Figure 2 Effects of stirring on the brownian bug clustering ($\Delta = 0.001$, $N_0 = 20,000$, as in Fig. 1). Each point is the position of a brownian bug in the (x, y) plane. **a**, Distribution at $t = 30\tau$ with $U/\tau/2 = 0.1$ and no birth or death. **b**, Distribution at $t = 1,000\tau$ and $U/\tau/2 = 0.1$.

hierarchy of distribution functions that describe the nonindependent probabilities of the positions of all brownian bugs²³. The second member of this hierarchy is the pair correlation function: $G(\mathbf{x}_1, \mathbf{x}_2, t)dA_1dA_2$ is the probability of finding a pair of brownian bugs with one member in the dA_1 surrounding \mathbf{x}_1 and the other member in the dA_2 surrounding \mathbf{x}_2 . For example, a Poisson point process with uniform concentration C_0 has $G = C_0^2$; departures of G from C_0^2 indicate pair correlations. In particular, because births occur at the same location as the parent, births increase the correlation G above C_0^2 for small $\mathbf{x}_1 - \mathbf{x}_2$. Because the pair correlation function G is the Fourier transform of the ‘coarse-grained’ concentration spectrum²⁴, we can relate this analysis to spectral descriptions of plankton patchiness^{23,9}.

For an isotropic and spatially homogeneous process, such as those in Figs 1 and 2, G depends only on the pair separation $r \equiv |\mathbf{x}_1 - \mathbf{x}_2|$. Then it is convenient to define the radial density function $g(r, t)$ by $G(\mathbf{x}_1, \mathbf{x}_2, t) = C_0^2 g(r, t)$ where C_0 is the uniform concentration. Because the pair correlations disappear at great distances, $g \rightarrow 1$ as $r \rightarrow \infty$. In a two-dimensional space, the expected number of brownian bugs located in the annulus $(r, r + dr)$ surrounding a typical brownian bug is $C_0 g(r, t) 2\pi r dr$; thus, by making a histogram of pair separations, we can estimate g from the results of a simulation (Fig. 3).

The solid curve in Fig. 3 shows that without random advection ($U = 0$) there is a greatly enhanced probability of finding one brownian bug in the close neighbourhood of another. The effect of random advection is to reduce g at $r = 0$ and to produce a long tail signifying slowly decaying pair correlations at large r (Fig. 3). The reduction at $r = 0$ quantifies the rapid advective separation of nearly coincident pairs, whereas the tail is the signature of the correlations that exist along the filaments in Fig. 2b. At moderate values of the stretching parameter U in equation (3), this tail is well described by $g \approx r^{-2}$ (Fig. 3): advection reduces reproductive pair correlations at small r , but enhances these correlations at large r .

The pair correlation function G is useful as a quantitative diagnostic of patchiness. But theoretical progress is also possible because for the brownian bug process $G(r, t)$ satisfies

$$G_t = 2Dr^{1-d}(r^{d-1}G_r)_r + 2(\lambda - \mu)G + \gamma r^{1-d}(r^{d+1}G_r)_r + 2\lambda C\delta(\mathbf{r}) \quad (2)$$

In this equation, d is the dimension of the space ($d = 2$ in all our simulations). The first term on the right-hand side is the diffusive separation of pairs with a diffusivity $2D$ because each member of the pair is an independent random walker. The second term on the right is the production of new pairs, which occurs if the birth rate λ exceeds the death rate μ . The third term on the right describes the separation of pairs by rapidly decorrelating random advection²¹, this particular form applies provided that r is much less than the smallest scale of the velocity field. The final term is production of coincident pairs (at $r = 0$) by birth.

If $\lambda = \mu$ then we can obtain a steady ($G_t = 0$) solution of equation (2). These are ‘constant-flux’ solutions: the reproductive source produces pairs at the origin ($r = 0$) of the pair-separation space. Then a combination of brownian motion and advective stretching transports these pairs outwards to larger values of r . Where r is much greater than $\sqrt{D/\gamma}$, advective stretching is the dominant process and the steady solution has $G \propto r^{-d}$ (this is the r^{-2} regime in Fig. 3). These steady analytic solutions of equation (2) are shown as dash-dot curves in Fig. 3. The large differences between theory and simulation at small r signifies the failure of the continuum approximation equation (2) once r is comparable to the step length of the random walk.

The brownian bug process is a simple model of plankton patchiness with two ingredients the production of concentration fluctuations by reproduction and the reduction in length scale of these fluctuations by stirring. It is crucial that the reproductive source term, $\delta(\mathbf{r})$ in equation (2), forces all wavenumbers equally. (We note

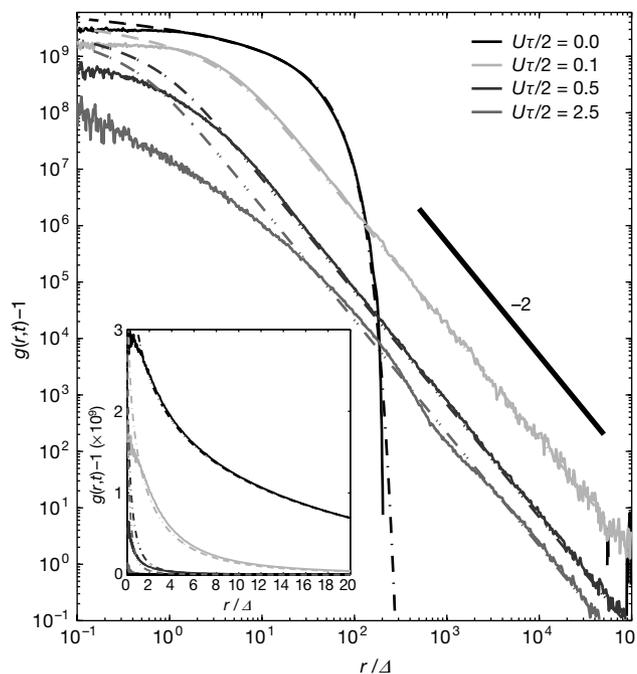


Figure 3 Logarithmic and linear plots of $g(r, t) - 1$ versus r/Δ . $t = 1,000\tau$, $N_0 = 20,000$, $\Delta = 10^{-7}$ and $U\tau/2 = 0, 0.1, 0.5, 2.5$. The dash-dot curves are

analytic solutions of equation (2). Inset, linear plot. The thick line labelled -2 indicates the r^{-2} scaling p [predicted by equation (2)].

that if one Fourier transform acts on equation (2), then $\delta(\mathbf{r}) \rightarrow 1$, showing that the final term forces a white spectrum.) Other models of plankton patchiness^{2,3,9,11} follow Batchelor²⁰ in assuming that concentration variations are forced externally only at small wavenumbers. The brownian bug process shows that the fluctuations can be generated by reproduction and that this process pumps variance into both large and small spatial scales. The common factor is that stirring reduces the length scale of concentration fluctuations so that diffusion can become effective. □

Methods

Conceptually the brownian bug model is a population of random-walking individuals evolving in continuous time and space and simultaneously undergoing a branching process²⁵. In probability theory this is known as a superbrownian process²⁶. To simulate this on a computer we must discretize.

The initial condition ($t = 0$) is prepared by randomly and independently placing $N_0 \gg 1$ brownian bugs (idealized as points) in the $L \times L$ square (periodically extended in order to avoid dealing with reflection boundary conditions). For visualization each brownian bug is tagged with a colour that varies smoothly with the y coordinate. The simulation is advanced through time in increments of a 'cycle' of duration τ . Each cycle consists of three steps: (1) random birth and death; (2) brownian motion; and (3) advective stirring. In step (1) each brownian bug reproduces by binary fission (with probability p) or dies (with probability q) or remains unchanged (with probability $1 - p - q$). When a lucky brownian bug divides, the offspring is placed on top of the parent and parents transmit their colour to their offspring. In step (2), bug k is displaced to a new position $\mathbf{x}'_k(t) = \mathbf{x}_k(t) + \delta\mathbf{x}_k(t)$. The components of $\delta\mathbf{x}_k$ are independent and identically distributed gaussian random variables, each with r.m.s. value Δ (that is, brownian motion with the diffusivity $D = \Delta^2/2\tau$). These independent displacements separate coincident parent-offspring pairs. In step (3) we use the random map procedure²⁴

$$x_k(t + \tau) = x'_k(t) + (U\tau/2) \cos[ky'_k(t) + \varphi(t)]$$

$$y_k(t + \tau) = y'_k(t) + (U\tau/2) \cos[kx'_k(t) + \theta(t)]$$

where $\varphi(t)$ and $\theta(t)$ are independent random phases and $k \equiv 2\pi/L$.

In principle, we can approach the continuous limit in equations (1) and (2) by taking $\tau \rightarrow 0$ while holding the parameters $\lambda \equiv p/\tau$, $\mu \equiv q/\tau$, $D \equiv \Delta^2/2\tau$ and $(kU)^2\tau$ fixed.

Step (1) is a Galton-Watson branching process²⁵ and because the probabilities of birth and death are equal (we take $p = q = 1/2$) we are at the critical point. In this case, even though the average population is always N_0 , the actual population has large fluctuations and total extinction is certain if the simulation lasts too long. We avoid this issue by making N_0 much larger than the number of generations. The suppression of extinction is a consequence of approaching the 'thermodynamic limit' ($N_0 \rightarrow \infty$ with $C_0 = N_0/L^2$ fixed).

The stretching parameter γ in equation (2) was determined for each value of $U\tau$ as the

best-fit slope of $(1/d)\langle \ln r(t) \rangle$ against $t \cdot \langle \ln r(t) \rangle$ is an average obtained from an ensemble of 800 particle pairs (without diffusion and birth/death) initially separated by $r(0) = 10^{-7}$. The growth rate λ was obtained by considering the case $U = 0$ and $\lambda = \mu$ and then least-squares-fitting the similarity solution of equation (2) to the results of a simulation in the range Δ to 200Δ . The resulting estimate of λ conforms with the expected result that $\lambda \approx 1/2\tau$.

Received 14 November 2000; accepted 4 June 2001.

1. Flierl, G., Grünbaum, D., Levin, S. & Olson, D. From individuals to aggregations: the interplay between behaviour and physics. *J. Theor. Biol.* **196**, 397-454 (1999).
2. Abraham, E. R. The generation of plankton patchiness by turbulent stirring. *Nature* **391**, 577-580 (1998).
3. Denman, K. L., Okubo, A. & Platt, T. The chlorophyll fluctuation spectrum in the sea. *Limnol. Oceanogr.* **22**, 1033-1038 (1977).
4. Franks, P. J. S. Spatial pattern in dense algal blooms. *Limnol. Oceanogr.* **42**, 1297-1305 (1997).
5. Kierstead, H. & Slobodkin, L. B. The size of water masses containing plankton blooms. *J. Mar. Res.* **12**, 141-147 (1953).
6. Neufeld, Z., López, C. & Haynes, P. H. Smooth-filament transition of active tracer fields stirred by chaotic advection. *Phys. Rev. Lett.* **82**, 2606-2609 (1999).
7. Robinson, A. R. On the theory of advective effects on biological dynamics in the sea. *Proc. R. Soc. Lond. A* **453**, 2295-2324 (1997).
8. Steele, J. H. & Henderson, E. W. A simple model for plankton patchiness. *J. Plankton Res.* **14**, 1397-1403 (1992).
9. Gower, J. F. R., Denman, K. L. & Holyer, R. J. Phytoplankton patchiness indicates the fluctuation spectrum of mesoscale oceanic structure. *Nature* **288**, 157-159 (1980).
10. Lesieur, M. & Sadourny, R. Satellite-sensed turbulent ocean structure. *Nature* **294**, 673 (1980).
11. Abraham, E. R. *et al.* Importance of stirring in the development of an iron-fertilized phytoplankton bloom. *Nature* **407**, 727-730 (2000).
12. Durrett, R. & Levin, S. The importance of being discrete (and spatial). *Theor. Pop. Biol.* **46**, 363-394 (1994).
13. Levin, S. & Segel, L. A. Hypothesis for origin of planktonic patchiness. *Nature* **259**, 659 (1976).
14. Segel, L. A. & Jackson, J. L. Dissipative structure: an explanation and an ecological example. *J. Theor. Biol.* **37**, 545-559 (1972).
15. Skellam, J. G. Random dispersion in theoretical populations. *Biometrika* **38**, 196-218 (1951).
16. Keeling, M. J., Mezić, I., Hendry, R. J., McGlade, J. & Rand, D. A. Characteristic length scales of spatial models in ecology via fluctuation analysis. *Phil. Trans. R. Soc. Lond. B* **352**, 1589-1601 (1997).
17. Pascual, M. & Levin, S. A. From individuals to population densities: searching for the intermediate scale of nontrivial determinism. *Ecology* **80**, 2225-2236 (1999).
18. Rand, D. A. & Wilson, H. B. Using spatio-temporal chaos and intermediate scale determinism in artificial ecologies to quantify spatially extended systems. *Proc. R. Soc. Lond. B* **259**, 55-63 (1995).
19. Batchelor, G. K. The effect of turbulence on material lines and surfaces. *Proc. R. Soc. Lond.* **213**, 349-366 (1952).
20. Batchelor, G. K. Small-scale variation of convected quantities like temperature in turbulent fluid. Part 1. General discussion and the case of small conductivity. *J. Fluid Mech.* **5**, 113-133 (1959).
21. Kraichnan, R. H. Convection of a passive scalar by a quasi-uniform random stretching field. *J. Fluid Mech.* **64**, 737-762 (1974).
22. Pierrehumbert, R. T. On tracer microstructure in the large-eddy dominated regime. *Chaos Solitons Fractals* **4**, 1091-1110 (1994).

23. van Kampen, N. G. *Stochastic Processes in Physics and Chemistry* Ch. 2 (North-Holland, Amsterdam, 1981).
24. Pierrehumbert, R. T. in *Nonlinear Phenomena in Atmospheric and Oceanic Sciences* Ch. 2 (Springer, New York, 1991).
25. Harris, T. E. *The Theory of Branching Processes* Ch. 1 (Springer, Berlin, 1963).
26. Le Gall, J. F. *Spatial Branching Processes, Random Snakes and Partial Differential Equations*, Lectures in Mathematics, ETH Zürich (Birkhäuser, Basel, 1999).

Acknowledgements

We thank E. Ben-Naim, R. Durrett, G. Flierl, P. Krapivsky, P. Morrison and F. Williams for discussion.

Correspondence and requests for materials should be addressed to W.R.Y. (e-mail: wryoung@ucsd.edu).

Evolution of digital organisms at high mutation rates leads to survival of the flattest

Claus O. Wilke*, Jia Lan Wang*, Charles Ofria†, Richard E. Lenski† & Christoph Adami*‡

* Digital Life Laboratory, Mail Code 136-93, California Institute of Technology, Pasadena, California 91125, USA

† Center for Biological Modeling, Michigan State University, East Lansing, Michigan 48824, USA

‡ Jet Propulsion Laboratory, Mail Code 126-347, California Institute of Technology, Pasadena, California 91109, USA

Darwinian evolution favours genotypes with high replication rates, a process called ‘survival of the fittest’. However, knowing the replication rate of each individual genotype may not suffice to predict the eventual survivor, even in an asexual population. According to quasi-species theory, selection favours the cloud of genotypes, interconnected by mutation, whose average replication rate is highest^{1–5}. Here we confirm this prediction using digital organisms that self-replicate, mutate and evolve^{6–9}. Forty pairs of populations were derived from 40 different ancestors in identical selective environments, except that one of each pair experienced a 4-fold higher mutation rate. In 12 cases, the dominant genotype that evolved at the lower mutation rate achieved a replication rate >1.5-fold faster than its counterpart. We allowed each of these disparate pairs to compete across a range of mutation rates. In each case, as mutation rate was increased, the outcome of competition switched to favour the genotype with the lower replication rate. These genotypes, although they occupied lower fitness peaks, were located in flatter regions of the fitness surface and were therefore more robust with respect to mutations.

Mutation and natural selection are the two most basic processes of evolution, yet the study of their interplay remains a challenging area for theoretical and empirical research. Recent studies have examined the effect of mutation rate on the speed of adaptive evolution^{10,11} and the role of selection in determining the mutation rate itself^{12–14}. Quasi-species models predict a particularly subtle interaction: mutation acts as a selective agent to shape the entire genome so that it is robust with respect to mutation^{1–5}. (See refs 15–18 for related predictions expressed in other terms.) In particular, selection in an asexual population should maximize the overall replication rate of a cloud of genotypes connected by mutation, rather than fix any one genotype that has the highest replication rate. Thus, a fast-replicating organism that occupies a high and narrow peak in the fitness landscape—where most nearby mutants

are unfit—can be displaced by an organism that occupies a lower but flatter peak. Thus, ‘survival of the flattest’ may be as important as ‘survival of the fittest’ at high mutation rates. This prediction has proved difficult to test experimentally, but a recent study¹⁹ with an RNA virus reported that two populations, derived from a common ancestor, have mutational neighbourhoods with different distributions of fitness effects.

Direct evidence for the displacement of a fast replicator by a more robust, slower one must come from experiments in which such organisms are squarely pitted against each other. The systematic (repeatable) winner of such a competition is, in effect, the fitter one, although the loser may have the higher replication rate. For example, imagine that a particular mutation yields a more robust genotype, but at the cost of a slightly lower replication rate. It is an empirical question whether the advantage of the mutational robustness is sufficient to offset its disadvantage in terms of replication rate. Quasi-species theory predicts that, under appropriate conditions (high mutation pressure), such a mutation can be fixed in an evolving population, despite its lower replication rate. This prediction does not depend on the details of the organism chosen for experiments, but only on mutation rate, replication speed, and robustness to mutations. Microorganisms, such as bacteria and viruses, are often used to test evolutionary theories, and competition experiments are typically performed to quantify fitness in the course of these tests. However, it would be difficult to disentangle the contributions of replication rate and robustness, because competitions measure the combined effect of both processes. Here, we use a more convenient system for disentangling these effects: digital organisms that live in, and adapt to, a virtual world created for them inside a computer.

Digital organisms are self-replicating computer programs that compete with one another for CPU (central processing unit) cycles, which are their limiting resource. Digital organisms have genomes (series of instructions) and phenotypes that are obtained by the execution of their genomic programs. The evolution of these programs is not simulated in the conventional (numerical) sense. Instead, they physically inhabit a reserved space in the computer’s memory (an ‘artificial Petri dish’), and they must copy their own genomes. Moreover, their evolution does not proceed towards a target specified in advance; instead it proceeds in an open-ended manner to produce phenotypes that are more successful in a particular environment. Digital organisms acquire resources (CPU cycles) by performing certain logical functions, much as biochemical organisms catalyse exothermic reactions to obtain energy. They lend themselves to evolutionary experiments because their environment can be readily manipulated to examine the importance of various selective pressures. The only environmental

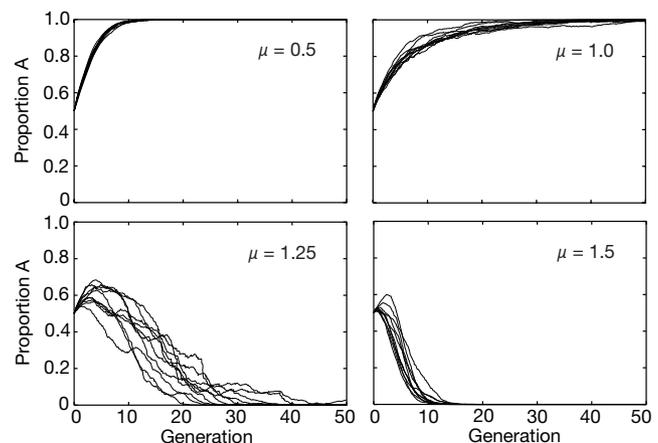


Figure 1 Competitions for one pair of organisms at four different mutation rates. Organism A replicates 1.96 times faster than B. μ, Genomic mutation rate.