

# Field parameterization and experimental test of the neutral theory of biodiversity

J. Timothy Wootton

Department of Ecology and Evolution, The University of Chicago, 1101 East 57th Street, Chicago, Illinois 60637, USA

Ecologists would like to explain general patterns observed across multi-species communities, such as species–area<sup>1</sup> and abundance–frequency relationships<sup>1–4</sup>, in terms of the fundamental processes of birth, death and migration underlying the dynamics of all constituent species. The unified neutral theory of biodiversity<sup>5,6</sup> and related theories<sup>7,8</sup> based on these fundamental population processes have successfully recreated general species–abundance patterns without accounting for either the variation among species and individuals or resource-releasing processes such as predation and disturbance, long emphasized in ecological theory<sup>9–14</sup>. If ecological communities can be described adequately without estimating variation in species and their interactions<sup>15</sup>, our understanding of ecological community organization and the predicted consequences of reduced biodiversity and environmental change would shift markedly. Here, I introduce a strong method to test the neutral theory that combines field parameterization of the underlying population dynamics with a field experiment, and apply it to a rocky intertidal community. Although the observed abundance–frequency distribution of the system follows that predicted by the neutral theory, the neutral theory predicts poorly the field experimental results, indicating an essential role for variation in species interactions.

Hubbell's neutral theory of biodiversity<sup>5,6</sup> (Box 1) views individuals in ecosystems as being partitioned into a series of communities of fixed size at a local scale, linked by migration from a regional pool of individuals of fixed population size. By fixing the size of the local and regional communities, the theory assumes that limiting resource levels are constant and are used in their entirety by the individuals in the system. The individuals in the regional pool are distributed across a set of ecologically equivalent species, and the pool is influenced minimally by any single local community. Random speciation events increase the regional species pool, and species extinction may occur both locally and regionally through a random-walk process that depends on species' population size. Because of the neutrality assumption, the regional pool converges to a defined distribution of individuals among species, controlled by a function, known as the fundamental biodiversity number  $\theta$ , which combines the regional population size and speciation rate. Local dynamics are typically scaled to a single death event and model the replacement of the dead individual as a function of the relative abundance of species in the regional pool, the relative species abundance in the local community, and probability of the new recruit being derived locally or via immigration from the regional pool.

As originally formulated, the neutral theory is accessible only to fairly weak empirical tests. Several of the key parameters underlying the model dynamics (regional population size, speciation rate, migration rate and death rate) are difficult or nearly impossible to measure directly. Hence, previous analyses have been limited to examining the consistency of the abundance–frequency distributions of local communities with distributions predicted by the neutral model. Such tests are weak, however, because both the observed and predicted data are ranked by abundance—guaranteeing that both observed and predicted functions decline monotonically—and because the predicted distributions are tuned by the  $\theta$

and migration rate parameters, permitting the predicted distributions to match a wide range of functional shapes<sup>16</sup>. A stronger test would be to generate estimates of model parameters from system dynamics, then test the predictions of the parameterized model in an independent situation that varied in known ways<sup>17</sup>, such as in a field experiment.

## Box 1 Original neutral model and an empirically accessible, resource-based reformulation.

The neutral theory<sup>5,6</sup> models the population dynamics at both local and regional scales, linked by recruitment from regional to local populations. The theory assumes ecologically identical individuals and complete utilization of limiting resources, resulting in approximately constant total numbers of individuals at both local and regional scales. Given these assumptions, the change in the size of a species' population at a local scale, modelled on the timescale of a single individual death, is

$$P_r(N_i - 1|N_i) = (N_i/J)[m(1 - p_i) + (1 - m)(\{J - N_i\}/\{J - 1\})]$$

$$P_r(N_i|N_i) = (N_i/J)[p_i m + (1 - m)(\{N_i - 1\}/\{J - 1\})] + (\{J - N_i\}/J) \times [m(1 - p_i) + (1 - m)(\{J - N_i - 1\}/\{J - 1\})] \quad (1)$$

$$P_r(N_i + 1|N_i) = (\{J - N_i\}/J)[p_i m + (1 - m)(N_i/\{J - 1\})]$$

where  $N_i$  is the local population size of species  $i$ ,  $J$  is the total local population size,  $m$  is the probability that the individual replacing the dead individual is derived from the regional, rather than local, species pool, and  $p_i$  is the proportion of the regional species pool comprised of species  $i$ . The theory models the dynamics of the regional pool similarly, but random speciation events, rather than migration from an external source, add new species. The regional pool will converge to a known species abundance distribution, with a shape that depends on a parameter  $\theta$  (referred to as the fundamental biodiversity number), which is a function of the total number of individuals contained in the regional pool and the per-individual speciation rate<sup>5,6</sup>.

The local dynamical equations of the neutral theory can be reformulated by changing perspective from individuals in the community to units of limiting resource used by individuals. Empirically, assigning transitions in resource (for example, space) users is often easier than determining which particular individual replaced an individual that just died, and facilitates relaxing the assumption of saturated resource use. A resource-based reformulation of the local dynamics of the neutral theory is

$$P_r(i_t|i_0) = (1 - D) + D[p_i m + (L - m)N_{i,0}/J] \quad (2)$$

$$P_r(j_t|i_0) = D[p_j m + (L - m)N_{j,0}/J]$$

$$P_r(i_t|0) = p_i m + (L - m)N_{i,0}/J$$

where  $P_r(j_t|i_0)$  is the probability that a unit of resource used by an individual of species  $i$  at an initial observation is used by species  $j$  at a subsequent observation  $t$  time units later,  $P_r(i_t|0)$  is the probability that a unit of unused resource is subsequently used by an individual of species  $i$ ,  $D$  is the probability that an individual using a resource unit does not survive over time  $t$ , and  $N_{i,0}$  is the population size of species  $i$  in the local community at the initial observation point. A new parameter  $L$  ( $m \leq L \leq 1$ ) allows exploration of recruitment limitation. When  $L = 1$ , no recruitment limitation to the resource occurs, as the neutral model assumes;  $L < 1$  implies recruitment limitation.

Combining equation (2) and expanding across all resource units describes the dynamics of the entire local population of each species  $i$

$$N_{i,t} = (1 - D)N_{i,0} + D[N_{i,0}(p_i m + (L - m)N_{i,0}/J) + (J - N_{i,0} - N_{0,0}) \times (p_i m + (L - m)N_{i,0}/J)] + N_{0,0}[p_i m + (L - m)N_{i,0}/J] \quad (3)$$

where  $N_{0,0}$  is the amount of unused resource in the system. Parameters in these equations can be estimated with empirical data on species abundances and transition probabilities in resource use without knowing a priori the individual death rate, assuming that  $L = 1$ , or estimating  $\theta$ .

By reformulating the theory to focus on resource units, particularly space or space-based resources, rather than individuals, many of the difficulties of parameter estimation can be overcome, permitting a strong test of the theory (Box 1). Specifically, by examining patterns of transitions among species using a known quantity of resource, death rate, migration rate and regional species composition can be estimated from local community dynamics, allowing parameterization of the local model without estimating  $\theta$ , and allowing experimental tests of the model in systems where limiting resources can be readily identified. Additionally, this perspective permits evaluation of the consequences of relaxing the assumption of saturated resource use (Box 1).

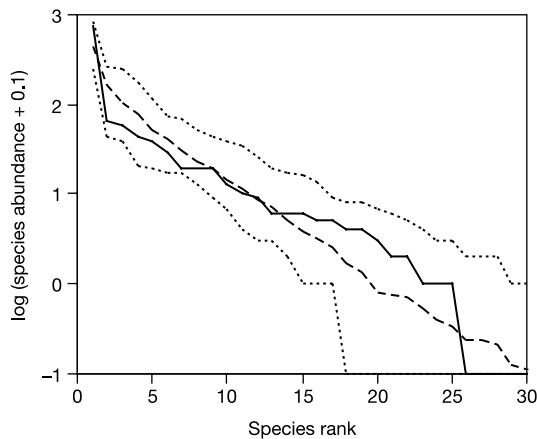
Hard-substrate marine benthic communities are ideal for testing the neutral theory for several reasons. First, they are characterized by an easily quantified limiting resource: attachment space on rocks<sup>17–19</sup>. Second, they exhibit relatively rapid dynamics, which allows relevant data on population dynamics to be collected over a reasonable timescale. Third, because of the moderate time and size scales over which the resident species operate, these systems have proven to be very amenable to field experimentation. Fourth, as predicted by the neutral model, they are well characterized as local communities linked to a regional source of individuals via oceanic dispersal. Finally, although studies in these systems have traditionally emphasized variation in species interactions<sup>13,18</sup>, the distribution of abundance across species follows that predicted by the neutral model (Fig. 1), suggesting a possible paradox.

I assigned parameters to the local dynamics equations of the neutral theory using 15,200 transitions from a 12-yr record of species occupancy of points in space on intertidal rock benches on Tatoosh Island, Washington, USA (see Methods). Given the sampling design, a local community is defined in this study as the 100 individuals occupying points in a spatially fixed 60 × 60 cm quadrat. The transitions (Fig. 1) were fit to the reformulated neutral model (Box 1) with maximum likelihood methods to generate the best estimates of parameters at the local population scale, using the species proportions observed in independent transects (Fig. 1) to guide the search of parameter space (see Methods). Maximum likelihood parameterization of the transition data in unmanipulated plots generated estimates for death rate ( $D$ ) and migration rate ( $m$ ) of 0.25 and 0.68, respectively, which are biologically reasonable given the dispersal syndromes and lifespans typical of the sessile

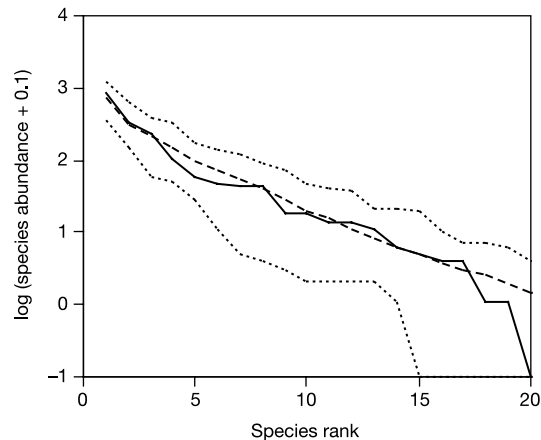
species in this system. The migration rate estimates generated independently from the analysis of transition data and from fitting the model to the abundance–frequency distribution of the community (0.71, Fig. 1) are quite similar, further suggesting that the neutral model might apply.

To test the predictions of the parameterized neutral model, I performed a parallel experiment in which I chronically deleted the dominant species in the system, *Mytilus californianus*, from nine 60 × 60 cm plots, by hand-removing any individuals when they entered the plots over the 7–11-yr duration of the experiment (see Methods). By preventing the establishment of *M. californianus* into plots, I experimentally removed this species from both the local and regional pool, thereby shifting the regional pool composition parameters in a known way by proportionally increasing the remaining species (see Methods). I then compared the predicted abundance of the remaining species under the assumption of neutral dynamics with the results of the species deletion experiment.

Fitting the general shape of the abundance–frequency distribution to data derived from experimental plots again was consistent with the neutral model, suggesting that it might be applicable (Fig. 2). Because the underlying model was parameterized, however, I could make specific predictions about which species should be associated with each abundance value. When I compared experimental results to these more refined predictions, the neutral model was not supported (Fig. 3). The neutral model predictions differed significantly compared with the range of species compositions observed in the experimental plots (randomization tests of multivariate distances,  $P < 0.001$ , see Methods), and explained –29.6% of the variance in mean species abundance. Hence, the model predictions actually added error over an assumption that there was no variation in species abundance. One possible explanation for the poor fit of the neutral model might be the neutral model's assumption that all resources are used. I explored this possibility by using the reformulated model (Box 1) and introducing a recruitment limitation parameter ( $L$ ) that was free to vary. This analysis yielded a best-fitting model with  $m = 0.66$ ,  $D = 0.10$  and  $L = 0.66$ . The equivalence of  $L$  and  $m$  implies that the local population contributed minimally to recruitment, which would not be surprising given the life histories of the sessile species in this system relative to the spatial scale of the local population. The predictions with these parameters, however, performed no better than the original neutral model ( $P < 0.001$ ), explaining –30.3% of the variance in



**Figure 1** Comparison of species abundance distribution from randomly placed transects in the middle intertidal zone of Tatoosh Island, Washington, USA, compared to the predicted abundance distribution under the neutral model (with 95% confidence interval). Observed (solid line), mean predicted (dashed line) and 95% confidence interval predicted (dotted lines) curves are shown. Best estimates of parameters:  $\theta = 5.7$ ,  $m = 0.71$ ,  $J = 1,128$  (parameter definitions in Box 1). Each species is arrayed along the x axis, ranked by abundance, and its abundance, on a log scale, is presented along the y axis.

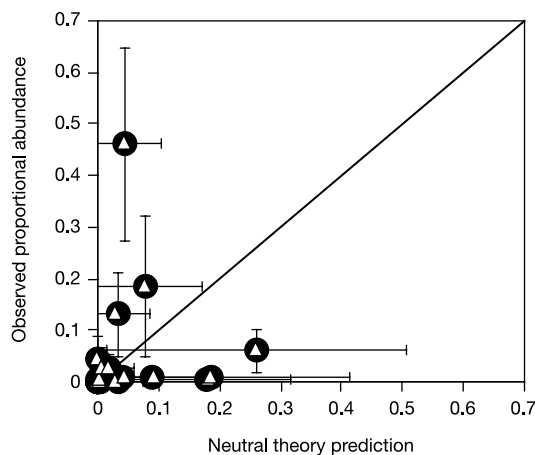


**Figure 2** Ranked abundance frequency distribution of individuals from last two censuses in experimental plots compared to the predicted abundance distribution under the neutral model (with 95% confidence interval) made without reference to species identity. Observed (solid line), mean predicted (dashed line) and 95% confidence interval predicted (dotted lines) curves are shown. Best estimates of parameters:  $\theta = 5.0$ ,  $m = 0.80$ ,  $J = 1,800$  (parameter definitions in Box 1).

mean species abundance in the experimental plots. Hence, the assumption of resource saturation could not explain the poor fit.

My results demonstrate that static abundance distribution patterns are a poor guide to identifying processes important to the organization of ecological communities, and illustrate the power of combining field experimentation with a focus on the population dynamic equations incorporating the essential processes of the neutral model, rather than on aggregate patterns several steps removed from ecological process. The results also clearly demonstrate that this community is not neutrally organized. The poor performance of the neutral model contrasts strongly with a Markov chain model incorporating species-specific differences<sup>20</sup>, which successfully predicted 84% of the variation in mean species abundance from mussel removal experiments<sup>21</sup>. Parameter variation in the Markov chain model could arise from differences in local and regional population sizes of species, recruitment limitation, or differences in species traits and interspecific interactions. Because only the last factor was absent from the neutral model and extensions that I explored, species differences seem to be essential elements in generating the structure and dynamics of this intertidal community. Hence, accurate prediction of the system response to external impacts requires an accounting of variation in species interactions, and is not primarily determined by global abundance patterns. This conclusion does not preclude essential elements of the neutral model—including metacommunity structure, dispersal dynamics and speciation–extinction processes—from having important roles in community dynamics. For example, changes in the abundance patterns observed with increasing spatial scale of the local community<sup>4,6</sup> might still be explained by sampling processes in a metacommunity context. Therefore, extending the theoretical framework of the neutral theory to include species differences may be rewarding.

If important assumptions of the neutral theory do not hold, why might patterns of species abundance follow a general pattern? One possibility is that species interaction strengths seem to follow a general pattern of distribution<sup>22–25</sup>, which may constrain species abundance patterns. Initial theoretical work suggests that such interaction strength distributions may arise from community



**Figure 3** Comparison between species-specific predictions from the parameterized neutral model (filled circles) or model with recruitment limitation (open triangles) and observed proportional abundances in experimental mussel removal plots (error bars: 95% confidence intervals). Variations in mean predicted abundance in the neutral model arise from differences in species abundance generated by random-walk processes at both local and regional scales. Error bars for predictions arise from varying initial conditions in different experimental plots, and from stochastic effects on small local populations under multiple model realizations. The diagonal line is the expected relationship for perfect fit of experimental data to model predictions.

assembly processes coupled with system stability and persistence<sup>26,27</sup>. Investigations of these issues and links to macroecological pattern are still in their infancy, but appear profitable to pursue both empirically and theoretically. □

**Methods**

Details of data collection and experimental manipulation can be found elsewhere<sup>20,21</sup>. Briefly, in 1993 I established 14 permanent 60 × 60 cm plots with 100 fixed points spaced regularly in a grid, and recorded the species present at each point in each plot annually in the late spring 1993–2004. I also collected independent data to characterize more broadly the species composition of the community by placing random transects through the mid-intertidal zone and identifying the sessile species present at each 5-cm interval along the transects<sup>20</sup>.

Following prior studies, I fit the neutral model to static species abundance distributions derived from the independent transect data<sup>20</sup> following the algorithm outlined by Hubbell<sup>6</sup>, which was implemented in a program written with MatLab software. Briefly, this algorithm assembled a neutral community of the size of the data sample (the local community size *J*) by randomly assigning individuals to a species based on their current abundance in the community, the probability of a new recruit migrating from the regional pool (*m*), and the probability that a migrant is a new species, which is governed by the fundamental biodiversity number  $\theta$ . I generated abundance distributions for a wide range of parameter values (0–1 for *m*, 1–20 for  $\theta$ ) and selected the parameter set that had the best fit to the distribution derived from the data, based on the maximum likelihood estimator<sup>6</sup>

$$L_e = \sum_{i=1,s} (N_i(\ln(N'_i/N_i) - (N'_i - N_i)))$$

where  $N'_i$  is the predicted abundance of species *i* for the sample size of the transect data,  $N_i$  is the actual abundance of species *i* in the transect data, and *s* is the number of species across all communities.

The local dynamic equations of the neutral model were fit using MatLab software to data from the annually censused plots using transitions in species occupancy at each point from one year to the next. One plot was not located in one year, reducing the number of transitions available by 200. The neutral theory was parameterized using these data by systematically evaluating parameter space for *p*, *m*, *D* and *L* in equation (3) (Box 1) and comparing the model fit to the data across all 14 plots and all 11 transition years using the maximum likelihood estimator

$$L_e = \sum_{t=1,11} \sum_{i=1,s} \sum_{j=1,14} (N_{i,j,t}(\ln(N'_{i,j,t}/N_{i,j,t}) - (N'_{i,j,t} - N_{i,j,t})))$$

where  $N'_{i,j,t}$  is the predicted abundance of species *i* in local plot *j* at time *t*, and  $N_{i,j,t}$  is the actual abundance of species *i* in plot *j* at time *t*. I estimated proportional species abundance of the regional pool ( $p_i$ ) using the neutral theory to generate 1,000 distributions of species abundance, basing these distributions on the best estimate of  $\theta$  derived from the species abundance distribution of 1,128 individuals from random transect data<sup>20</sup> (Fig. 1), and assuming a regional population of 10,000 individuals. Convergence of runs of the neutral model with 10,000 individuals and with larger population sizes is generally very high<sup>6</sup>. To maximize search efficiency, I assigned species to particular abundances generated by the neutral theory based on their abundance rank in the transect data<sup>20</sup>, because the proportional species abundance of a large sample of individuals reflects the regional pool under the neutral model<sup>6</sup>. I also used the proportional abundance of species in the transect data as one possible estimate of the regional distribution. For each distribution of proportional abundance, I evaluated model fit while systematically varying *m*, *D* and *L* in 0.01-unit increments across the range of possible parameter values (0–1). The parameter combination that maximized the likelihood estimator was considered the proper description of neutral model dynamics. When zeros occurred in data or model predictions, I added 0.5 to both observed and predicted abundance to avoid problems of taking logs of zero.

I predicted the community response to deletion of *M. californianus* under the neutral model by setting  $p_{Mc} = 0$  and adjusting the rest of the regional pool proportionately to sum to 1 (that is,  $p'_i = p_i / (1 - p_{Mc})$ , where  $p'_i$  is the adjusted proportion of species *i* in the regional pool after *M. californianus* (subscript Mc) is removed). I then simulated the neutral model for experimental plots of 100 individuals using the observed initial conditions and the duration of each of the nine replicates. Five other replicates<sup>20</sup> were not used because only data for per cent cover, not number of individuals at defined census points, were available and because the neutral theory predictions depend on knowing local population sizes. Analyses including these replicates and assuming a population size of 100 gave similar results. Each plot was simulated ten times under the neutral model, and the mean composition of these simulations was taken as the model prediction for each year. I compared the composition of individuals from the final two years of each experimental replicate and model simulation using a multivariate randomization approach. Specifically, I asked whether the mean of the absolute (euclidean) distance in multivariate space (that is,  $[\sum_{i=1,2} (N_{i,1} - N_{i,2})^2]^{0.5}$ , where  $N_{i,1}$  and  $N_{i,2}$  are the number of individuals of species *i* in community 1 and 2, respectively) between model predictions and experimental results for each of the nine plots tended to be higher than the mean euclidean distance between nine randomly chosen pairs of experimental plots. Higher distances between model predictions and experimental results would indicate a poorer fit of the model to the data than expected by chance. This comparison was repeated 1,000 times, and the fraction of times that model predictions had lower distances than experimental pairs was taken as the probability that the model predictions did not differ from the observed data. Other metrics (variance predicted, Bray–Curtis distance) were also examined to verify the robustness of the analysis, and did not change the conclusions.

Received 24 September; accepted 16 November 2004; doi:10.1038/nature03211.

1. MacArthur, R. H. & Wilson, E. O. *The Theory of Island Biogeography* (Princeton Univ. Press, Princeton, 1969).
2. Fisher, R. A., Corbet, A. S. & Williams, C. B. The relation between the number of species and the number of individuals in a random sample of an animal population. *J. Anim. Ecol.* **12**, 42–58 (1943).
3. Preston, F. W. The commonness, and rarity, of species. *Ecology* **41**, 611–627 (1948).
4. Brown, J. H. *Macroecology* (Univ. Chicago Press, Chicago, 1995).
5. Hubbell, S. P. A unified theory of biogeography and relative species abundance and its application to tropical rain forests and coral reefs. *Coral Reefs* **16**, S9–S21 (1997).
6. Hubbell, S. P. *The Unified Theory of Biodiversity and Biogeography* (Princeton Univ. Press, Princeton, 2001).
7. Caswell, H. Community structure: a neutral model analysis. *Ecol. Monogr.* **46**, 327–354 (1976).
8. Bell, G. Neutral macroecology. *Science* **293**, 2413–2418 (2001).
9. Elton, C. S. *Animal Ecology* (Sidgwick and Jackson, London, 1927).
10. Gause, G. F. *The Struggle for Existence* (Hafner, New York, 1934).
11. Hutchinson, G. E. Homage to Santa Rosalia or why are there so many kinds of animals? *Am. Nat.* **93**, 145–159 (1959).
12. Huffaker, C. B. Experimental studies on predation: dispersion factors and predator-prey oscillations. *Hilgardia* **27**, 343–383 (1958).
13. Paine, R. T. Food web complexity and species diversity. *Am. Nat.* **100**, 65–75 (1966).
14. MacArthur, R. H. *Geographical Ecology* (Harper & Row, New York, 1972).
15. Laska, M. S. & Wootton, J. T. Theoretical concepts and empirical approaches to measuring interaction strength. *Ecology* **79**, 461–476 (1998).
16. McGill, B. J. Strong and weak tests of macroecological theory. *Oikos* **102**, 679–685 (2003).
17. Adler, P. B. Neutral models fail to reproduce observed species-area and species-time relationships in Kansas grasslands. *Ecology* **85**, 1265–1272 (2004).
18. Connell, J. H. The influence of interspecific competition and other factors on the distribution of the barnacle *Chthamalus stellatus*. *Ecology* **42**, 710–723 (1961).
19. Paine, R. T. Ecological determinism in the competition for space. *Ecology* **65**, 1339–1348 (1984).
20. Wootton, J. T. Prediction in complex communities: analysis of empirically-derived Markov models. *Ecology* **82**, 580–598 (2001).
21. Wootton, J. T. Markov chain models predict the consequences of experimental extinctions. *Ecol. Lett.* **7**, 653–660 (2004).
22. Paine, R. T. Food-web analysis through field measurements of per capita interaction strength. *Nature* **355**, 73–75 (1992).
23. Moore, J. C., de Ruiter, P. C. & Hunt, H. W. The influence of productivity on the stability of real and model ecosystems. *Science* **261**, 906–908 (1993).
24. Raffaelli, D. G. & Hall, S. J. in *Food Webs: Integration of Pattern and Dynamics* (eds Polis, G. & Winemiller, K.) 185–191 (Chapman and Hall, New York, 1996).
25. Wootton, J. T. Estimates and tests of per-capita interaction strength: diet, abundance, and impact of intertidally foraging birds. *Ecol. Monogr.* **67**, 45–64 (1997).
26. Kokkoris, G. D., Troumbis, A. Y. & Lawton, J. H. Patterns of species interaction strength in assembled theoretical competition communities. *Ecol. Lett.* **2**, 70–74 (1999).
27. Drossel, B., McKane, A. & Quince, C. The impact of nonlinear functional responses on the long-term evolution of food web structure. *J. Theor. Biol.* **229**, 539–548 (2004).

**Acknowledgements** I thank the Makah Tribal Council for providing access to Tatoosh Island; J. Sheridan, J. Salamunovitch, F. Stevens, A. Miller, B. Scott, J. Chase, J. Shurin, K. Rose, L. Weis, R. Kordas, K. Edwards, M. Novak, J. Duke, J. Orcutt, K. Barnes, C. Neufeld and L. Weintraub for field assistance; and NSF, EPA (CISES) and the Andrew W. Mellon foundation for partial financial support.

**Competing interests statement** The author declares that he has no competing financial interests.

**Correspondence** and requests for materials should be addressed to J.T.W. (twoodton@uchicago.edu).

## Evolutionary dynamics on graphs

Erez Lieberman<sup>1,2</sup>, Christoph Hauert<sup>1,3</sup> & Martin A. Nowak<sup>1</sup>

<sup>1</sup>Program for Evolutionary Dynamics, Departments of Organismic and Evolutionary Biology, Mathematics, and Applied Mathematics, Harvard University, Cambridge, Massachusetts 02138, USA

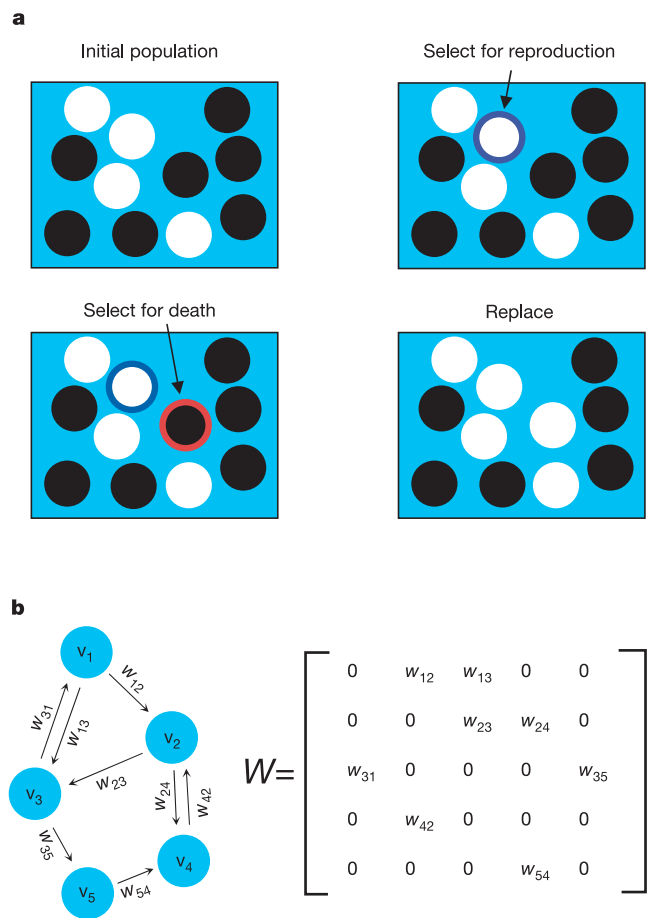
<sup>2</sup>Harvard-MIT Division of Health Sciences and Technology, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

<sup>3</sup>Department of Zoology, University of British Columbia, Vancouver, British Columbia V6T 1Z4, Canada

Evolutionary dynamics have been traditionally studied in the context of homogeneous or spatially extended populations<sup>1–4</sup>. Here we generalize population structure by arranging individuals on a graph. Each vertex represents an individual. The weighted edges denote reproductive rates which govern how

often individuals place offspring into adjacent vertices. The homogeneous population, described by the Moran process<sup>5</sup>, is the special case of a fully connected graph with evenly weighted edges. Spatial structures are described by graphs where vertices are connected with their nearest neighbours. We also explore evolution on random and scale-free networks<sup>5–7</sup>. We determine the fixation probability of mutants, and characterize those graphs for which fixation behaviour is identical to that of a homogeneous population<sup>7</sup>. Furthermore, some graphs act as suppressors and others as amplifiers of selection. It is even possible to find graphs that guarantee the fixation of any advantageous mutant. We also study frequency-dependent selection and show that the outcome of evolutionary games can depend entirely on the structure of the underlying graph. Evolutionary graph theory has many fascinating applications ranging from ecology to multi-cellular organization and economics.

Evolutionary dynamics act on populations. Neither genes, nor cells, nor individuals evolve; only populations evolve. In small populations, random drift dominates, whereas large populations



**Figure 1** Models of evolution. **a**, The Moran process describes stochastic evolution of a finite population of constant size. In each time step, an individual is chosen for reproduction with a probability proportional to its fitness; a second individual is chosen for death. The offspring of the first individual replaces the second. **b**, In the setting of evolutionary graph theory, individuals occupy the vertices of a graph. In each time step, an individual is selected with a probability proportional to its fitness; the weights of the outgoing edges determine the probabilities that the corresponding neighbour will be replaced by the offspring. The process is described by a stochastic matrix  $W$ , where  $w_{ij}$  denotes the probability that an offspring of individual  $i$  will replace individual  $j$ . In a more general setting, at each time step, an edge  $ij$  is selected with a probability proportional to its weight and the fitness of the individual at its tail. The Moran process is the special case of a complete graph with identical weights.