

# Artificial ecosystem selection

William Swenson, David Sloan Wilson\*, and Roberta Elias

Department of Biological Sciences, Binghamton University, Binghamton, NY 13902-6000

Communicated by Lynn Margulis, University of Massachusetts, Amherst, MA, May 23, 2000 (received for review November 8, 1999)

**Artificial selection has been practiced for centuries to shape the properties of individual organisms, providing Darwin with a powerful argument for his theory of natural selection. We show that the properties of whole ecosystems can also be shaped by artificial selection procedures. Ecosystems initiated in the laboratory vary phenotypically and a proportion of the variation is heritable, despite the fact that the ecosystems initially are composed of thousands of species and millions of individuals. Artificial ecosystem selection can be used for practical purposes, illustrates an important role for complex interactions in evolution, and challenges a widespread belief that selection is most effective at lower levels of the biological hierarchy.**

The question of whether selection can operate above the level of the individual has been controversial among evolutionary biologists (1). Artificial selection experiments allow this question to be asked in a straightforward manner: A population of units (individuals, groups, communities, or ecosystems) is established in the laboratory, the units are measured with respect to a phenotypic trait, and some units are selected on the basis of this trait to serve as “parents” for creating the next generation of units. If phenotypic variation has a heritable basis, then the average value of the trait for the offspring generation will shift in the direction of selection. Thus, the fundamental ingredients of phenotypic variation and heritability that are required for selection to produce evolutionary change can be experimentally examined at all levels of the biological hierarchy.

Artificial selection of single-species groups and two-species communities has demonstrated heritable variation at these levels (2). Indeed, the response to higher-level selection has been greater than expected on the basis of theoretical models, a point that we will return to below. Our experiments involve the selection of soil ecosystems for the phenotypic trait of above-ground plant biomass and aquatic ecosystems for the phenotypic trait of hydrogen ion concentration (pH). In an additional set of experiments reported elsewhere (W.S., J. Arendt, and D.S.W., unpublished results), we show that ecosystems can be selected for practical purposes, such as the ability to degrade toxic compounds. In all of our experiments, the ecosystems are laboratory microcosms that are physically small but large in biological terms, initially including thousands of species of bacteria, algae, protozoa, fungi, and other small organisms numbering many millions of individuals. An ecosystem is defined as the interactions of species with each other and their physical environment. Because two of the phenotypic traits (pH and degradation of the toxic compound 3-chloroaniline) are properties of the physical environment and the third trait (above-ground plant biomass) is likely to be mediated through effects on the physical environment, our experiments qualify as selection at the ecosystem level.

## Materials and Methods

All selection experiments require a population of units, a phenotypic trait that forms the basis of selection, and a method for creating a new generation of units from the selected units of the previous generation. In the soil ecosystem experiments, the units were transparent containers designed to grow plants under microbially controlled conditions (GA-7 acrylic vessels and couplers; Magenta Corporation, Chicago). Each container was

filled with 85 g of well-mixed potting soil that was dried and rewetted with 25 ml of water and sterilized by autoclaving within the containers. Approximately 50 surface-sterilized seeds of *Arabidopsis thaliana* (strain Landsberg erecta) were added to each container, and the plants were allowed to grow under continuous light at 25°C for 35 days, which constituted a single ecosystem generation. The phenotypic trait of above-ground plant biomass was measured for each container by clipping the plants at the soil surface and drying to constant weight at 80°C before weighing.

At the beginning of the experiment, a sample of unsterilized soil from a single source (a hemlock-hardwoods forest located on the Binghamton University campus) was made into a slurry by adding sterilized water, and to provide the biotic component of the soil ecosystems, a controlled amount of soil was added to the sterile soil within the containers before adding plant seeds. During the course of the experiment, the soil from the selected units of each generation was used to inoculate the units of the next generation. The plants were not a part of the evolving ecosystems because seeds were added from an external source every generation. Plant biomass was merely the phenotypic trait of the soil ecosystem that was used as the basis of selection. Inoculation of microcosms, planting, and harvesting were carried out under sterile conditions so as to minimize contamination of the microcosms.

Two sets of high and low lines were initiated in which the size of the soil inoculum differed by two orders of magnitude (6.0 g vs. 0.06 g). Each line consisted of 15 units and the 3 units with the highest (or lowest) value of the phenotypic trait were used as parents by combining the soil from the 3 units into a slurry that was used to inoculate the “offspring” generation of units. Fifteen microcosms per generation for the high and low treatments of the 6.0-g and 0.06-g inoculum sizes and two control treatments led to a total of 90 microcosms per generation for the complete experiment, which was continued for 16 35-day ecosystem generations. In generations 1 through 9, we ran sterile control treatments to serve as baselines that consisted of high and low selection lines inoculated with autoclaved 6.0-g slurries. To test for physical effects of the slurry density, starting at generation 10 the control treatments consisted of 6.0 g vs. 0.06 g sterilized slurry lines in which the three parent ecosystems for the next generation were randomly chosen. Microcosms from all treatments were randomly interspersed and rearranged within the growth chamber three times per week to minimize environmental variation. To examine the effects of ecosystem selection on the physical properties of the soil, soil samples from generations 13 and 14 of the 6.0-g high and low lines were sent to the Cornell Agricultural and Life Sciences Soil Nutrient Analysis Laboratory (Ithaca, NY), where they were analyzed for  $\text{NO}_3^-$ ,  $\text{NH}_4^+$ , K, Ca, Mg, P, Fe, Mn, Zn, Cu, Al, and pH.

In the aquatic ecosystem experiment, approximately 2 ml of sediment and 28 ml of water from a pond on the campus of

\*To whom reprint requests should be addressed. E-mail: [dwilson@binghamton.edu](mailto:dwilson@binghamton.edu).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Article published online before print: *Proc. Natl. Acad. Sci. USA*, 10.1073/pnas.150237597. Article and publication date are at [www.pnas.org/cgi/doi/10.1073/pnas.150237597](http://www.pnas.org/cgi/doi/10.1073/pnas.150237597)

Binghamton University were placed in each of 72 test tubes, which were incubated for 3.5 days on a 14.5-h light cycle with “day” and “night” temperatures of 25°C and 20°C, respectively. Each test tube then was measured for pH, which served as the phenotypic trait for ecosystem-level selection. In one line, the six ecosystems with the highest pH were used as parents for the next generation of ecosystems. Each offspring ecosystem was composed of 5 ml from a single parent plus 25 ml of autoclaved water and sediment from the same pond. Thus, unlike the soil ecosystem experiments, the selected parent ecosystems were not mixed to create the offspring ecosystems. Four offspring were created from each parent for a total of 24 ecosystems per generation. In a second line, the six ecosystems with the lowest pH were used as parents in a similar fashion. Thereafter, each generation of the high and low line was derived from the six test tubes with the highest and lowest pH, respectively, from the previous generation. In a third line, six ecosystems were selected at random each generation to serve as parents, without regard to their pH. The three treatments led to a total of 72 test tubes per generation for the total experiment, which was continued for 40 3.5-day ecosystem generations.

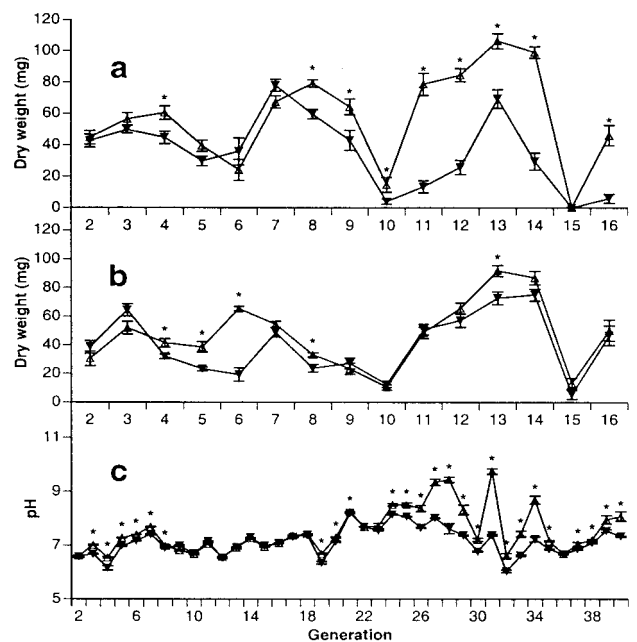
Because the phenotypic traits in these experiments were not necessarily normally distributed, tests of statistical differences between mean phenotypes were carried out by using bootstrapped comparisons, programmed specifically for the experiments, in which randomizations were iterated  $10^4$  times (standard parametric tests yielded nearly identical results). Soil nutrient data were subjected to a multivariate discriminant function analysis performed with STATISTICA release 4.1 for the Macintosh.  $\text{NO}_3^-$  and Mg were dropped from the analysis because of redundancy with other factors.

## Results

For ecosystem-level selection to work, (i) the ecosystems must vary in the phenotypic trait that is being selected, (ii) phenotypic variation must be caused by underlying properties of the ecosystems, and (iii) offspring ecosystems must partially resemble their parents in their effect on the phenotypic trait. The results of the experiments are shown in Fig. 1. The absolute values of the phenotypic traits varied considerably from generation to generation, which we attribute to subtle but important environmental factors that we were unable to hold constant. Despite this environmental source of phenotypic variation, which decreases heritability in artificial selection experiments, the high and low lines diverged from each other in two of the three experiments. Fig. 2 displays the results as deviations from the mean phenotypic value every generation, which removes the environmental “noise” and allows the patterns of divergence to be seen more clearly.

All three experiments exhibited an initial divergence between the high and low lines, followed by a collapse that lasted for a number of generations before a second and greater divergence in two of the three experiments. A fungus that appeared in both lines and killed virtually all of the plants caused the collapse of differences during generation 15 of the 6.0-g soil experiment. The fungus did not reappear in generation 16, and the difference between lines reinstated itself. Plant biomass in the high and low lines of the 6.0-g experiment differed by an average factor of 4.2 for generations 10–16 when generation 15 is excluded. pH in the high and low lines of the aquatic ecosystem experiment differed by an average factor of 1.1 for generations 25–40. Because pH is measured on a logarithmic scale, this represents a 25-fold difference in hydrogen ion concentration.

Through generation 9 of the soil experiments, the 6.0-g sterilized-inocula controls showed similar generation-to-generation variation as that seen in the experimental groups. However, differences between the high and low selection lines were not significant, and there was no evidence of a trend toward



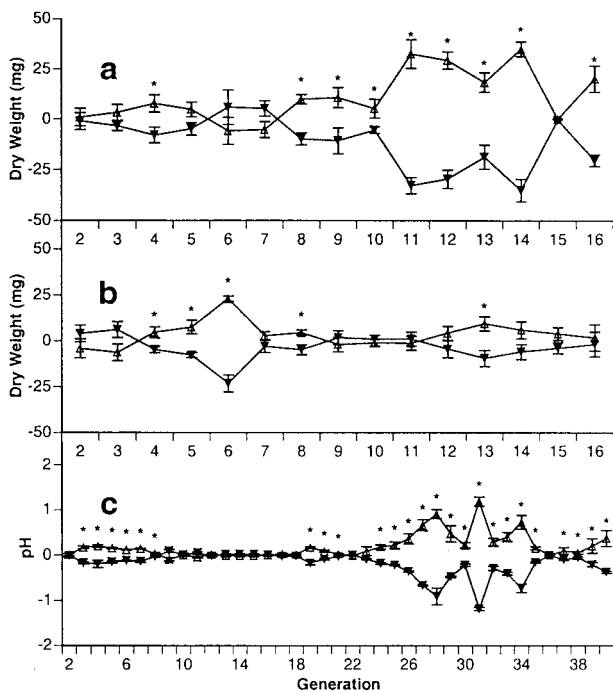
**Fig. 1.** Results of soil ecosystem and aquatic ecosystem selection experiments. (a) Above-ground biomass of *Arabidopsis thaliana* plants grown in ecosystems inoculated with 6.0 g of soil. Open, upward-pointing triangles represent mean of 15 microcosms ( $\pm 1$  SE) selected for high biomass. Solid, downward-pointing triangles represent selection for low biomass. Asterisks indicate generations in which biomass differed significantly (bootstrap test,  $P < 0.05$ ). There were no significant differences between mean plant biomass in either of the two paired control treatments in the soil ecosystem experiments. (b) The 0.06-g soil inoculum experiment, symbols as described in a. (c) pH of aquatic microcosms selected for high or low pH. Means of 24 aquatic microcosms ( $\pm 1$  SE), symbols analogous to those given in a.

divergence between the lines. Nor were there any significant differences or trends when comparing sterilized, nonselected 6.0-g inocula and 0.06-g inocula in generations 11–15 (no control treatments were run in generation 16). In the aquatic experiment, the pH of the nonselected control line closely resembled that in the line selected for greater acidity, indicating that the divergence between the experimental lines largely was attributable to ecosystem selection for decreased acidity.

Response to selection at the phenotypic level presumably requires an underlying change in ecosystem processes that cause the phenotypic traits in a proximate sense. We tested this proposition for the 6.0-g soil experiment by analyzing the soil of the control, high, and low lines at the time of harvest during generations 13 and 14. The discriminant function analysis (Fig. 3) clearly discriminates among the three lines on the basis of soil properties. Because the soil was identical at the beginning of each generation, these differences reflect a divergence in the biotic components of the ecosystems that change the soil chemistry over the course of a generation. The most important soil variables that discriminate between the high and low lines include potassium, zinc, and phosphorous content, which plausibly affect plant biomass. Detailed understanding of the ecosystem processes that proximally cause the selected phenotypic traits will require additional research.

## Discussion

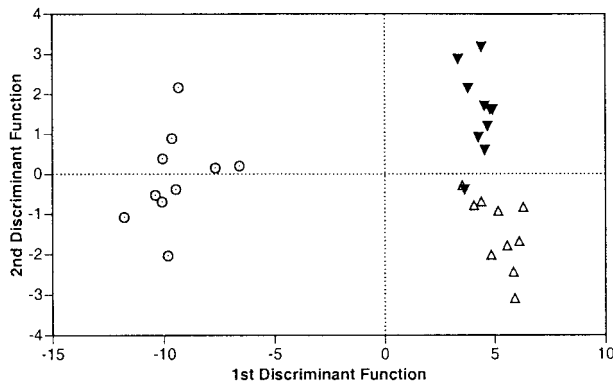
The response to ecosystem-level selection in our experiments stands in contrast to theoretical models that have led many to reject higher-level selection as an important evolutionary force (3). However, the divergence of the high and low lines in our experiments also has an unstable quality, initially collapsing in all



**Fig. 2.** Results of ecosystem artificial selection experiments (as in Fig. 1) expressed as deviations from overall means. Symbols are the same as in Fig. 1. (a) Difference between above-ground biomasses of *Arabidopsis thaliana* grown in ecosystem microcosms selected at high or low biomass, 6.0-g inoculum. (b) Difference between above-ground biomasses, 0.06-g inoculum treatment. (c) Difference in pH of aquatic microcosms selected for high or low pH.

three experiments and reappearing in only two. The 4.2-fold difference in plant biomass and the 25-fold difference in pH are impressive responses to ecosystem-level selection but they, too, may have collapsed if the experiments were continued for a longer period.

**Evolution and Complex Systems.** Both the response to selection and its unstable quality can probably be explained in terms of



**Fig. 3.** Discriminant function analysis of 10 soil nutrient variables in the 6.0-g inoculum size treatments, generations 13 and 14 combined. Open, upward-pointing triangles represent soils in microcosms selected for high biomass. Filled, downward-pointing triangles represent selection for low biomass. Circles represent soils from microcosms inoculated with autoclaved slurries and selected at random. The first discriminant function is largely attributable to soil  $\text{NH}_4^+$  content. The second discriminant function is most strongly attributable to soil K, Zn, and P content. (Wilk's  $\lambda$ , 0.00828;  $F(20, 156) = 77.895$ ;  $P < 0.0001$ ).

complex systems dynamics. Mathematical and computer simulation models of evolution tend to assume a simple relationship between phenotypic traits and their genetic basis, such as an altruistic behavior that is coded directly by an altruistic gene. In the case of our soil ecosystem experiment, we might imagine genes in soil organisms that have positive or negative effects on plant biomass, with variation among ecosystems caused by sampling error. Ecosystems initiated by small numbers of soil organisms may possess sufficient variation for ecosystem-level selection to operate, but not ecosystems initiated by large numbers of soil organisms. By this reasoning, our experiment should not have worked because variation caused by sampling error would be negligible in ecosystems initiated by 6.0 g of soil, which contains many millions of microbes comprising thousands of species (4–6). Also, if variation among ecosystems is caused purely by sampling error, it should have been greater in our 0.06-g experiments than our 6.0-g experiments, which differed by two orders of magnitude in the number of organisms initiating each ecosystem.

The substantial phenotypic variation observed in our experiments (and also in artificial selection experiments at the group and two-species community levels) reveals that something is wrong with theoretical reasoning based on simple interactions. In real biological systems, phenotypic traits often are caused by complex interactions among components of the system. Complex physical systems such as the weather are well known to display sensitive dependence on initial conditions, which causes replicate systems that initially are virtually identical to become very different over time (the so-called “butterfly effect”; ref. 7). If complex biological systems are like complex physical systems, they too will diverge in their phenotypic properties, no matter how small their initial differences. In the context of our experiment, very small initial differences in the genetic and species composition of our units, initially caused by sampling error, will develop into much larger differences during the course of an ecosystem generation, with correlated effects on the phenotypic trait that forms the basis for selection. Ecosystems initiated by 6.0 g of soil might vary as much as ecosystems initiated by 0.06 g of soil, because the initial differences that form the basis of the butterfly effect can be arbitrarily small. In short, theoretical reasoning based on complex interactions leads to profoundly different predictions about phenotypic variation among units, one of the fundamental ingredients of natural selection at all levels.

Complex interactions also are likely to have an effect on heritability that might explain the unusual patterns of response to selection shown in Fig. 2. Consider the replicate ecosystems at the beginning of an ecosystem-selection experiment. Sensitive dependence will cause them to diverge in their species composition and the genetic composition of the component species. Some of the differences will influence the phenotypic trait being measured. However, the genetic and species compositions of the selected ecosystems are unlikely to have come to equilibrium by the time they are used to create the next generation. The offspring ecosystems are therefore likely to embark on trajectories of their own with different effects on the phenotypic trait being measured, which will appear as an absence of heritability in the artificial selection experiment. Eventually, some ecosystems may reach a relatively stable local equilibrium with two properties: (i) they produce the phenotype that enables them to survive as units in the artificial selection experiment, and (ii) they are internally stable enough to retain their properties in the passage from the parent to the offspring generation, which appears as heritability in the artificial selection experiment. Artificial ecosystem selection can be seen as a method of searching an astronomically large parameter space (thousands of species and thousands of genes within each species) for ecosystems with these two properties.

**Research Design and Criticisms.** In designing our research, we decided to maximize the number of experiments at the expense of replication within each experiment. Thus, we included two ecosystems (soil and water), two phenotypic traits (plant biomass and pH), two inoculum sizes (6.0 g and 0.06 g), and two methods of forming offspring ecosystems (mixing the parent ecosystems vs. keeping them separate, roughly analogous to sexual and asexual reproduction). However, each experiment consisted of only a single high line, a single low line, and a sterile control line. Lack of replication means that we can be statistically confident that the high and low lines diverged in two of the three experiments, but we cannot say, for example, that ecosystem-level selection is more effective when 6.0 g rather than 0.06 g of soil is used as an inoculum. Replicated lines within each experiment would be required to make this and other statements.

This research strategy was justified by the fact that ecosystem-level selection had never been attempted before and was regarded by many of our colleagues as a risky venture. It therefore seemed wise to try a number of combinations of systems, traits, and other parameter values rather than placing all of our eggs in one basket with one well-replicated experiment. Now that we have demonstrated the basic possibility of ecosystem-level selection, it is appropriate to shift the research strategy in the direction of higher replication. Our newest set of experiments (W.S., J. Arendt, and D.S.W., unpublished results) selected aquatic ecosystems for the ability to degrade the environmental pollutant 3-chloroaniline in four replicated lines, as compared with four lines in which parent ecosystems were randomly chosen without respect to their degradation ability. The eight lines initially diverged widely in their degradation abilities, providing another demonstration of the butterfly effect. Three of the four selected lines increased their degradation ability during the course of 30 ecosystem generations, as compared with none of the four nonselected lines. These results, along with the results of the experiments reported here, provide solid evidence that heritable phenotypic variation can exist at the level of laboratory ecosystems, allowing their properties to be shaped by artificial selection experiments.

Thinking of ecosystems as units of selection is sufficiently unique that we would like to discuss a number of misplaced criticisms that we have encountered in the past, which can be dispelled by comparing our experiments with more familiar artificial selection experiments at the individual level. First, the phenotypic traits of plant biomass, pH, and chloroaniline degradation seem arbitrary and do not measure fundamental ecosystem processes. The same “criticism” could be applied with equal force to Mendel’s choice of flower color and the wrinkled vs. smooth texture of peas as individual-level phenotypic traits. Artificial selection experiments have always been somewhat arbitrary in their choice of phenotypic traits (e.g., the fanciful plumage of chickens) or have selected traits that are useful for humans without being “fundamental” as far as the organism is concerned. Furthermore, it is difficult to define what counts as fundamental from an evolutionary perspective. If a seemingly superficial trait such as coat color enhances survival and reproduction, why should it be considered less fundamental than another trait such as oxygen transport? In the future it will be interesting to know whether more fundamental properties of ecosystems, however defined, will respond to ecosystem selection, but in general the concept of phenotype needs to be as broad at the ecosystem level as it is at the individual level.

A second criticism is that we do not know the species composition (or the genetic composition within species) of the ecosystems that we have selected. If we haven’t identified the actual strains of microbes and how they interact with each other, isn’t our research sloppy? Again, most artificial selection experiments at the individual level are conducted entirely in terms of phenotypes without detailed knowledge of the genetic and

physiological mechanisms that underlie the response to selection. Details of the phenotypic response to selection often are used to infer underlying mechanisms (e.g., particulate inheritance in the case of Mendel, or additive vs. epistatic genetic interactions in the case of modern quantitative genetics research), but this is also true in the case of our experiments, which point to complex interactions as a source of phenotypic variation and heritability. Mechanistic knowledge is always desirable but is not required to conduct an artificial selection experiment. Indeed, it is a strength of evolutionary thinking in general that it can proceed so far in the absence of mechanistic understanding. Darwin’s theory of natural selection was developed in complete ignorance of the mechanistic basis of inheritance. Following Darwin, it is a legitimate research strategy to first demonstrate the existence of heritable phenotypic variation at the ecosystem level, and only then to attempt to understand its mechanistic basis.

A third criticism is that our results can be explained without invoking ecosystem-level selection. It may be convenient to call a trait such as plant biomass or pH an ecosystem-level trait, but, ultimately, it is caused by species and their genes. Similarly, even though we selected whole ecosystems on the basis of their traits, ultimately the experimental procedure selected certain species and genes. Why, then, cannot the results be interpreted as a form of species-level or gene-level selection? This argument reflects a widespread misunderstanding about multilevel selection theory, which has always assumed a genetic basis for the traits that evolve (1). For example, a typical group selection model attempts to show how a gene for altruism can evolve by increasing the fitness of whole groups, despite being selected against within groups. For the model to work, there must be a process of natural selection at the group level (a population of groups, variation among groups, etc.) that counterbalances natural selection within groups. Even when group selection succeeds, however, it is still a gene (for altruism) that evolves and replaces an alternative gene (for selfishness) in the global population. To say that the altruistic gene is “selfish” in some ultimate sense, simply because it evolves, is not an argument against group selection. In the parlance of selfish gene theory (8, 9), this misunderstanding confuses the concept of genes as “replicators,” with the concept of “vehicles” of selection, which are the engines of evolutionary change and can exist any level of the biological hierarchy. In our experiments, species and genes evolve by causing some ecosystems to be selected over others. The ecosystem is clearly the vehicle of selection, despite the fact that species and genes qualify as replicators.

**Natural Ecosystem Selection?** Just as Darwin used artificial individual selection to argue for an analogous natural process, it is reasonable to postulate a process of natural ecosystem selection. Ecosystems have traditionally been envisioned at a large spatial scale, such as a forest, a lake, or even the whole earth (10, 11). Natural ecosystem selection is unlikely to occur at these scales and is theoretically impossible for the whole earth, unless we are willing to speculate about between-planet selection. Nevertheless, by miniaturizing the concept of ecosystems, it becomes possible to envision a process of natural ecosystem selection very similar to our experiments. Potential examples include endo- and ectosymbiotic associations (12–17) and selection among the “microecosystems” of bacteria and algae that form on organic aggregates and account for most of the productivity in marine and aquatic environments (18–20). These miniature ecosystems are still large in biological terms (many species and many more individuals), but thousands or even billions of them exist within a forest or a lake, allowing the differential survival and productivity of ecosystems to become an important part of the evolutionary process. In addition, groups, communities, and ecosystems do not require discrete boundaries to be units of selection.

The essential ingredient is localized interactions, such that one patch fares better than another on the basis of its properties, even when the boundaries between patches are fuzzy. Thus, plant–soil and plant–plant interactions, which are localized but usually lack discrete boundaries, still may be subject to ecosystem-level selection. Properly conceived, it is possible to imagine a natural searching process for adaptive and internally stable ecosystems that has been occurring among innumerable units for innumerable generations.

**Using Artificial Ecosystem Selection for Practical Purposes.** Regardless of whether ecosystem selection occurs in nature, it still can be used to evolve “designer ecosystems” in the laboratory for practical purposes. The most difficult problems in human life cannot be solved by single individuals and require coordinated teams of specialists. Difficult biological problems, such as the breakdown of toxic substances in the soil, might similarly require coordinated teams of species rather than a single species. Creating these teams from the “bottom up,” by testing many different species in many combinations, is possible in principle but difficult in practice. Ecosystem selection provides a simple “top down” alternative, by creating a large number of ecosystems

and selecting those that best solve the problem to create the next generation of ecosystems. Even if the ultimate goal is a reductionistic understanding of the “team” members and how they work together, ecosystem selection provides a much more effective screening method than does testing all possible members in all possible combinations. Our experiments on chloroaniline degradation (W.S., J. Arendt, and D.S.W., unpublished results) are a first step in this direction.

Artificial ecosystem selection is simple and logistically feasible. It has many practical applications and encourages us to look for a similar process of natural ecosystem selection. The main reason that it has not been attempted before seems to be conceptual. Perhaps because it was regarded as theoretically implausible, it simply didn’t occur to biologists to look for heritable phenotypic variation at the level of ecosystems, which allows their properties to be shaped by standard artificial selection procedures.

This work was supported by National Science Foundation Grant DEB-9707650 to D.S.W. and a Howard Hughes Medical Institute Undergraduate Biological Sciences grant to R.E.

1. Sober, E. & Wilson, D. S. (1998) *Unto Others: The Evolution of Altruism* (Harvard Univ. Press, Cambridge, MA).
2. Goodnight, C. J. & Stevens, L. (1997) *Am. Nat.* **150**, S59–S79.
3. Williams, G. C. (1966) *Adaptation and Natural Selection: A Critique of Some Current Evolutionary Thought* (Princeton Univ. Press, Princeton).
4. Torsvik, V., Goksøyr, J. & Daae, F. L. (1990) *Appl. Environ. Microbiol.* **56**, 782–787.
5. Borneman, J., Skroch, P. W., O’Sullivan, K. M., Palus, J. A., Rujmanek, N. G., Jansen, J. L., Nienhuis, J. & Triplett, E. W. (1996) *Appl. Environ. Microbiol.* **62**, 1935–1943.
6. Bintrim, S. B., Donohue, T. J., Handelsman, J., Roberts, G. P & Goodman, R. M. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 277–282.
7. Lorenz, E. N., ed. (1993) in *The Essence of Chaos* (Univ. of Washington Press, Seattle), pp. 181–184.
8. Dawkins, R. (1976) *The Selfish Gene* (Oxford Univ. Press, New York).
9. Dawkins, R. (1982) *The Extended Phenotype: The Long Reach of the Gene* (Oxford Univ. Press, New York).
10. Lovelock, J. E. & Margulis, L. (1974) *Tellus* **XXVI**, 2–9.
11. Lenton, T. M. (1998) *Nature (London)* **394**, 439–447.
12. Wilson, D. S. (1983) *Am. Nat.* **121**, 851–870.
13. Wilson, D. S. & Knollenberg, W. G. (1987) *Evol. Ecol.* **1**, 139–159.
14. Frank, S. A. (1994) *J. Theor. Biol.* **170**, 393–400.
15. Frank, S. A. (1996) *Am. Nat.* **148**, 1113–1124.
16. Frank, S. A. (1996) *Proc. R. Soc. London B* **263**, 339–344.
17. Margulis, L. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 1071–1076.
18. Goldman, J. C. (1984) *Bull. Mar. Sci.* **35**, 462–476.
19. Stone, L. & Weisburd, R. S. J. (1992) *Trends Ecol. Evol.* **7**, 263–267.
20. Stone, L. & Berman, T. (1993) *Bull. Math. Biol.* **55**, 919–936.