# Adaptive Radiation from Resource Competition in Digital Organisms

Stephanie S. Chow,[1]* Claus O. Wilke,[1,2]* Charles Ofria,[3] Richard E. Lenski,[4] Christoph Adami[1,2]†

**Species richness often peaks at intermediate productivity and decreases as resources become more or less abundant. The mechanisms that produce this pattern are not completely known, but several previous studies have suggested environmental heterogeneity as a cause. In experiments with evolving digital organisms and populations of fixed size, maximum species richness emerges at intermediate productivity, even in a spatially homogeneous environment, owing to frequency-dependent selection to exploit an influx of mixed resources. A diverse pool of limiting resources is sufficient to cause adaptive radiation, which is manifest by the origin and maintenance of phenotypically and phylogenetically distinct groups of organisms.**

The explanation for differences in species richness among habitats has been called "perhaps the greatest unsolved ecological riddle" (*1*). One factor widely thought to control species richness is productivity. Productivity can be defined in several ways, including as "the rate at which energy flows through an ecosystem" (*2*). Species richness typically grows with productivity but sometimes declines at very high productivity levels (*3*–*9*). Several alternative hypotheses have been proposed for why diversity might be maximized at intermediate productivity, with this relation depending in various ways on spatial heterogeneity in resource distribution or on predation that prevents competitive exclusion (*1*–*5*, *9*–*13*).

Adaptive radiation has been directly observed in bacteria, whose rapid generations allow experiments on an evolutionary time scale. For *Escherichia coli* grown in a homogeneous medium with a single limiting sugar, adaptive radiation can occur via cross feeding, in which metabolites produced by certain genotypes serve as resources for others (*14*–*18*). On the other hand, spatial heterogeneity was necessary for adaptive radiation in *Pseudomonas fluorescens*, because populations evolving in a spatially homogeneous environment showed no diversification (*9*, *19*). A Levene-type model was proposed to explain this diversification under heterogeneous, but not homogeneous, conditions (*9*). However, adaptive radiation should also occur in spatially uniform environments that contain multiple resources (*20*, *21*), because negative frequency-dependent selection, in which competition favors rare types, is a general mechanism for maintaining diversity.

Our focus here, as in the experiments with bacteria, is on adaptive radiation giving rise to species that are functionally distinct from one another; our focus is not on processes that can lead to reproductive isolation, because bacteria and digital organisms reproduce asexually. Thus, one can think of a species as a group of organisms that share most of their genotypic and phenotypic features but which differ substantially from members of other species (*22*). Digital organisms allow us to quantify precisely the extent of phylogenetic divergence, a metric relevant to both asexual and sexual species. Specifically, we examine the relation between the productivity of an environment and the species richness that arises via evolution from a single ancestral type.

Digital organisms are self-replicating computer programs that inhabit a computational environment in which they compete for resources, mutate, and evolve (*23*, *24*). They can be viewed as domesticated computer viruses, and they have been used to study a number of fundamental problems in evolution (*23*, *25*). Each organism has a genome that is a sequence of instructions in a special-purpose programming language; these instructions allow an organism with a suitable genotype to copy its genome and perform various logic functions on a virtual CPU (central processing unit). The logic functions correspond to exothermic catalytic reactions insofar as they provide the organism with energy that allows faster reproduction. Mutations arise during the copy process and affect single instructions at the rate of 0.5 mutations per genome replication in this study. Most mutations are deleterious, often resulting in an organism unable to reproduce, but some are neutral or beneficial. Beneficial mutations may result from refining the replication mechanism, improving an existing logic function, or discovering a new function that is rewarded (*25*).

In previous work with digital organisms, the resources corresponding to different logic functions were unlimited in the sense that the reward obtained for performing any particular function was unchanged as the number of organisms that performed it increased (although the relative advantage of performing a function declined as more organisms performed it). In this study, we used a configuration in which the reward obtained by a particular organism for performing any logic function declines with consumption of the reward by other organisms (*26*). Thus, each logic function corresponds to a distinct resource that can be independently depleted and renewed (*27*). The environment is configured as a well-stirred reactor without any spatial structure. Resources flow into the reactor at a constant rate, and they flow out in proportion to their abundance. In our experiments, inflow rates are varied over six orders of magnitude, whereas the outflow rate is kept constant (*27*). Population size is held constant at 3000 organisms across all inflow rates; newborn organisms replace randomly chosen existing organisms. We inoculate the populations with a simple organism that receives some basal energy and can reproduce but which cannot perform any logic functions. Nine different resources corresponding to nine different logic functions (*25*) are provided, all at the same inflow rate in a given experiment. There are no predators or parasites on the digital organisms in our experiments. Thus, only the overall productivity of the environment varies, with the relative supply rate of different resources and all other factors held constant. As our operational definition of species, we use clusters of organisms that all have small phylogenetic distances from one another (*27*). The phylogenetic distance between two organisms is defined as the total number of intermediate organisms (having different genotypes from their parents) along the lines of descent leading to their most recent common ancestor (*25*).

Productivity, as defined here by resource inflow to the system, has a profound effect on the resulting species richness. Figure 1 shows the number of species (determined by the clustering algorithm) as a function of resource inflow rate and time. Species richness reaches its maximum level at intermediate productivity and declines substantially at both lower and higher productivities (Fig. 1A). Species richness also increases with time but approaches a plateau rather than increasing indefinitely (Fig. 1B). At very low productivity, the reward for performing new logic functions is too low to offset the cost of

[1]Digital Life Laboratory 136-93, California Institute of Technology, Pasadena, CA 91125, USA. [2]Keck Graduate Institute, 535 Watson Drive, Claremont, CA 91711, USA. [3]Department of Computer Science and Engineering, [4]Department of Microbiology and Molecular Genetics, Michigan State University, East Lansing, MI 48824, USA.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: adami@caltech.edu

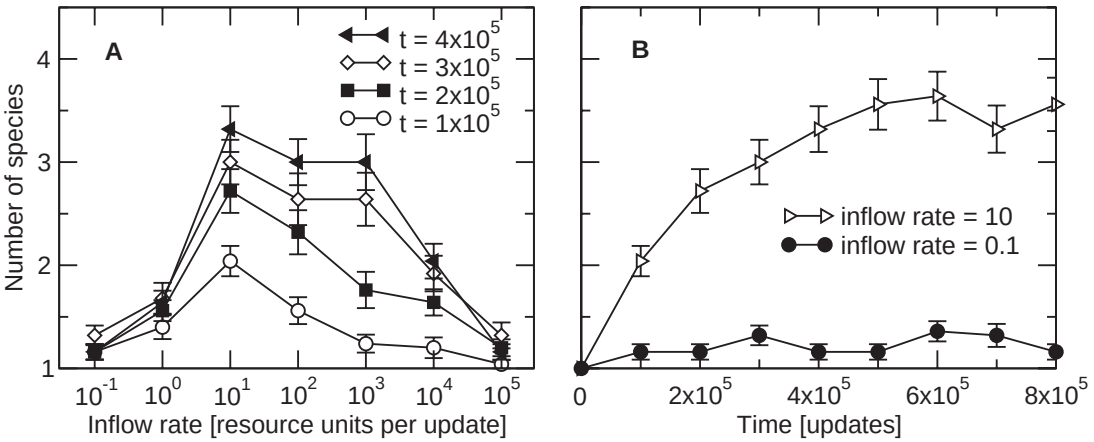


**Fig. 1.** Mean number of species as a function of inflow rate (**A**) and time (**B**). Species are defined with the use of a clustering algorithm based on phylogenetic distance. Each point represents 25 replicate runs, with bars showing standard errors. All of the runs start with a simple ancestor that is unable to use any of the resources obtained by performing logic functions.

performing them; the evolved organisms that dominate continue to use only the basal energy that was also available to their ancestor. At intermediate productivity, species richness increases over time as organisms with new logic functions evolve and invade the community (Fig. 1B). No single species dominates the community owing to the frequency-dependent advantage associated with the use of abundant resources that have been newly discovered or that remain underutilized. The resulting community is stable in the absence of mutation and further evolution, and species can invade when rare (Fig. 2 and fig. S1). At higher productivity levels, the logic functions that enable various resources to be used are potentially advantageous, but frequency-dependent selection becomes weaker as inflows increase and the depletable resources become effectively unlimited (fig. S2). At the highest productivity level, the evolved community is dominated by a single specialist species that can grow the fastest. Because the resources associated with the simplest logic functions are sufficiently abundant to fuel the whole population, these organisms do not carry out the more costly logic functions, nor can specialists on the more costly functions gain a foothold.

Figure 3 shows phylogenetic depth versus time for a typical experiment performed at an intermediate productivity level. The adaptive radiation begins quite early, and the resulting species differences are permanent. Of course, many other branches appear and disappear; the figure shows only those lineages that have descendants at the end of the experiment. By contrast, no permanent branches corresponding to stable species are seen at the lowest and highest productivity levels (fig. S3). Notice, too, that each branch displays a fairly constant rate of genetic change, a sort of molecular clock (*28*, *29*), although there is variation in rates among the species. All species derive from the same ancestor, and they all experience the same per-instruction mutation rate. The rate differences between species must reflect differences in the proportion of conserved sites that evolve in the different species' genomes, because total genome length is fixed at 100 in these experiments. Other replicates are qualitatively similar. Among the replicates at each inflow rate, there is of course variation in the resulting number of species that evolve. The maximum final species number that we have observed at any productivity level is six. Occasionally, a new species evolves that can

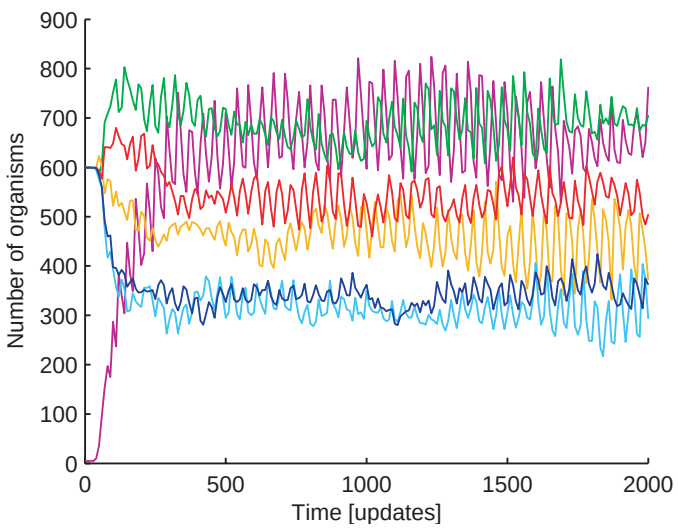


**Fig. 2.** Invasion when rare with numbers of organisms shown as a function of time. Colors correspond to the same six species as in Fig. 3. Species 1 in this example ecosystem can invade when rare, and all species quickly approach an equilibrium abundance (inflow rate is $10^3$ resource units per update). Invasions of the other species are shown in fig. S1.

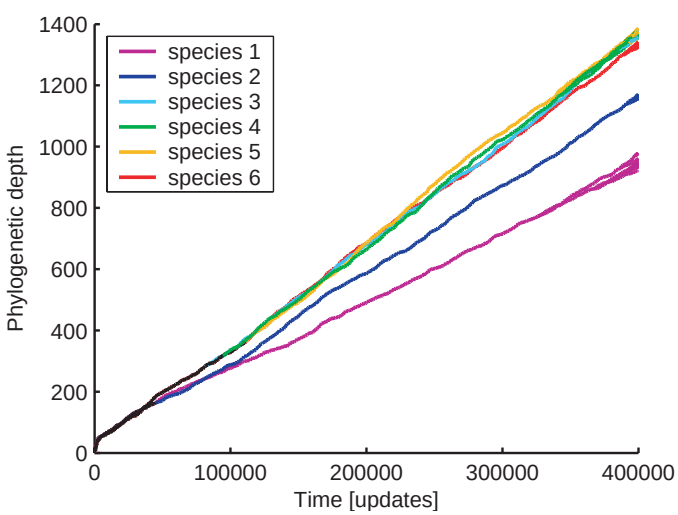


**Fig. 3.** Phylogenetic depth profile of a typical experiment at an inflow rate of $10^3$ resource units per update as a function of time. The lines trace the number of evolutionary steps (generations in which parent and offspring differ by one or more new mutations) leading from each organism in the final community all the way back to the common ancestor. Each color corresponds to a different species in the final population as determined by the clustering algorithm. The fact that two branches may remain close to one another or even cross does not imply that their relative genetic distance is small, because the figure merely shows the one-dimensional projection of total genetic change from the ancestral reference point. Typical phylogenetic depth profiles at lower and higher inflow rates are shown in fig. S3.

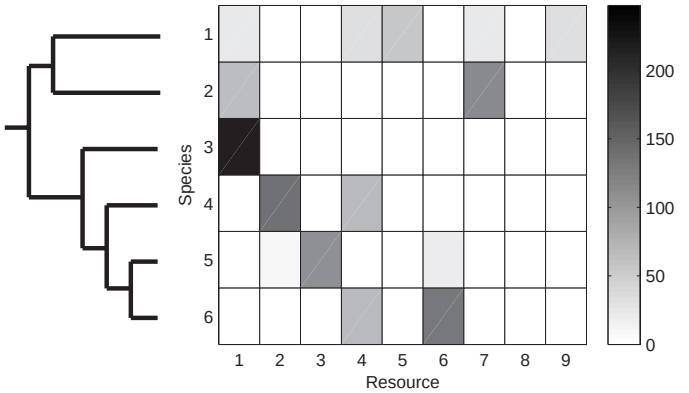


**Fig. 4.** Matrix of resource usage by the species shown in Fig. 3. Each entry is shaded according to the number of times a particular resource is used during the life cycle of an average member of the species. Although there is some overlap in resource use, each species dominates on at least one resource. The branching pattern shown to the left of the usage matrix indicates the relation between the species based on the phylogeny in Fig. 3 (relative branch lengths not to scale).

simultaneously outcompete two existing species, thereby reducing the total number of species in the community, although species richness often subsequently rebounds. Thus, the species richness in these digital communities reflects a dynamic steady state rather than some optimal end point.

Moving from a phylogenetic perspective to one based on ecological performance, Fig. 4 displays the resource utilization patterns of the same community of species illustrated by the phylogenetic depth profile in Fig. 3. Resource usage levels are averaged over all organisms that belong to a given species based on the clustering algorithm, which uses only phylogenetic data (not ecological phenotypes). The six species are all clearly phenotypically distinct. Moreover, each species consumes a larger share of at least one resource than does any other species. Species 3 is specialized on the use of resource 1, whereas the other species use two or more resources. Species 1 has the most generalized phenotype, using five resources including two that it dominates (resources 5 and 9). The precise partitioning of resources varies from one replicate experiment to another and does not follow any obvious rule. We have also demonstrated that our speciation results are not dependent on the particular ancestral organism we have used and that the pattern of species richness is not substantially altered when species that evolved at other resource levels are introduced into a community [Supporting Online Material (SOM) Text].

In summary, our experiments demonstrate an evolutionary effect of productivity on species richness, with adaptive radiation leading to maximum diversity at intermediate resource inflows in communities with fixed total population size. The decline in species richness at high productivity occurs because selection shifts from favoring exploitation of unused resources to favoring maximum replication rate when resources are superabundant. Importantly, this pattern does not require spatial heterogeneity or predation, although these and other ecological factors may augment or oppose the adaptive radiation.

### References and Notes


1. P. J. Morin, *Nature* **406**, 463 (2000).
2. M. L. Rosenzweig, Z. Abramsky, in *Species Diversity in Ecological Communities*, R. E. Ricklefs, D. Schluter, Eds. (Univ. Chicago Press, Chicago, IL, 1993), pp. 52–56.
3. D. Tilman, S. Pacala, in *Species Diversity in Ecological Communities*, R. E. Ricklefs, D. Schluter, Eds. (Univ. Chicago Press, Chicago, IL, 1993), pp. 13–25.
4. M. L. Rosenzweig, *Species Diversity in Space and Time* (Cambridge Univ. Press, Cambridge, 1995).
5. R. B. Waide *et al*., *Annu. Rev. Ecol. Syst.* **30**, 257 (1999).
6. G. G. Mittelbach *et al*., *Ecology* **82**, 2381 (2001).
7. T. Fukami, P. J. Morin, *Nature* **424**, 423 (2003).
8. M. C. Horner-Devine, M. A. Leibold, V. H. Smith, B. J. M. Bohannan, *Ecol. Lett.* **6**, 613 (2003).
9. R. Kassen, A. Buckling, G. Bell, P. B. Rainey, *Nature* **406**, 508 (2000).
10. D. Tilman, *Resource Competition and Community Structure* (Princeton Univ. Press, Princeton, NJ, 1982).
11. R. D. Holt, J. Grover, D. Tilman, *Am. Nat.* **144**, 741 (1994).
12. P. A. Abrams, *Ecology* **76**, 2019 (1995).
13. M. A. Leibold, *Am. Nat.* **147**, 784 (1996).
14. R. B. Helling, C. N. Vargas, J. Adams, *Genetics* **116**, 349 (1987).
15. R. F. Rosenzweig, R. R. Sharp, D. S. Treves, J. Adams, *Genetics* **137**, 903 (1994).
16. P. E. Turner, V. Souza, R. E. Lenski, *Ecology* **77**, 2119 (1996).
17. D. S. Treves, S. Manning, J. Adams, *Mol. Biol. Evol.* **15**, 789 (1998).
18. D. E. Rozen, R. E. Lenski, *Am. Nat.* **155**, 24 (2000).
19. P. B. Rainey, M. Travisano, *Nature* **394**, 69 (1998).
20. M. Doebeli, U. Dieckmann, *Am. Nat.* **156**, S77 (2000).
21. M. L. Friesen, G. Saxer, M. Travisano, M. Doebeli, *Evolution* **58**, 245 (2004).
22. F. M. Cohan, *Syst. Biol.* **50**, 513 (2001).
23. C. O. Wilke, C. Adami, *Trends Ecol. Evol.* **17**, 528 (2002).
24. C. Ofria, C. O. Wilke, *Artif. Life* **10**, 191 (2004).
25. R. E. Lenski, C. Ofria, R. T. Pennock, C. Adami, *Nature* **423**, 139 (2003).
26. T. F. Cooper, C. Ofria, in *Artificial Life VIII: Proceedings of the Eighth International Conference on Artificial Life*, International Society for Artificial Life, Sydney, Australia, 9 to 13 December 2002, R. Standish, M. A. Bedau, H. A. Abbass, Eds. (MIT Press, Cambridge, MA, 2003), p. 227.
27. Materials and methods are available as supporting material on *Science* Online.
28. E. Zuckerkandl, L. Pauling, in *Evolving Genes and Proteins*, V. Bryson, H. J. Vogel, Eds. (Academic Press, New York, 1965), pp. 97–166.
29. M. Kimura, *The Neutral Theory of Molecular Evolution* (Cambridge Univ. Press, Cambridge, 1983).


30. Supported by grant DEB-9981397 from NSF.


**Supporting Online Material**

www.sciencemag.org/cgi/content/full/305/5680/84/DC1
Materials and Methods
SOM Text

2 February 2004; accepted 3 June 2004

# Self-Assembling Protein Microarrays

**Niroshan Ramachandran,[1] Eugenie Hainsworth,[1,2] Bhupinder Bhullar,[1] Samuel Eisenstein,[1] Benjamin Rosen,[1] Albert Y. Lau,[1] Johannes C. Walter,[3] Joshua LaBaer[1]***

**Protein microarrays provide a powerful tool for the study of protein function. However, they are not widely used, in part because of the challenges in producing proteins to spot on the arrays. We generated protein microarrays by printing complementary DNAs onto glass slides and then translating target proteins with mammalian reticulocyte lysate. Epitope tags fused to the proteins allowed them to be immobilized in situ. This obviated the need to purify proteins, avoided protein stability problems during storage, and captured sufficient protein for functional studies. We used the technology to map pairwise interactions among 29 human DNA replication initiation proteins, recapitulate the regulation of Cdt1 binding to select replication proteins, and map its geminin-binding domain.**

To exploit the growing number of expression-ready cDNA clone collections, high-throughput (HT) methods to study protein function are needed (*1*–*5*). The development of protein microarrays offers one compelling approach (*6*–*8*). Protein microarrays are currently available in two general formats. Antibody arrays contain an array of antibodies that measure the abundance of specific proteins (or other molecules) in samples (*9*). Our work focuses on target protein arrays, which present arrayed proteins of interest. They can be used to examine target protein interactions with other molecules, such as drugs, antibodies, nucleic acids, lipids, or other proteins. In addition, the arrays can be interrogated to find substrates for enzymes (*10*, *11*).