

The Effects of Dispersal on the Evolution of Artificial Parasites

Carlo C. Maley
Room E105
Department of Zoology
South Parks Road
Oxford, OX1 3PS
ENGLAND¹

email: cmaley@vax.oxford.ac.uk
or maley@occs.cs.oberlin.edu
phone: 44-865-792328

Ray's Tierra model was extended to investigate the effects of distance of dispersal on the evolution of parasitism. Tierra's environment was enriched from one to two dimensions. Within this two-dimensional environment, artificial life evolved under one of two conditions. Under the *adjacent dispersal* condition, an artificial organism's offspring were placed next to itself in the environment. In contrast, under the *global dispersal* condition, an organism's offspring were placed randomly in the environment. The *adjacent dispersal* condition provided a more hostile environment for the evolution of parasites than the *global dispersal* condition. Specifically, it took significantly longer for parasites to invade and dominate the population under the local dispersal condition. There was also an increase in genetic variation under the *adjacent dispersal* condition, which is probably due to the emergence of a patchy environment and the subsequent local adaptations to that environment.

¹After July, 1993, I can be reached at 545 Technology Sq. NE43-418, MIT, Cambridge, MA, 02139, USA.

Background

Spatial Structure

Models that focus on spatial structure have had much success in the past decade. Cellular Automata (CA) are the pre-eminent form of spatially structured models and have been used for studying physical (Bak et al. 1988), chemical (Dewdney 1988) and biological phenomena (Langton 1986). In particular, the addition of spatial structure, via a CA, to the hypercycle theory of the origin of life has recently revitalized that theory (Dyson 1985). The hypercycle theory has been rejected in part because hypercycles are not evolutionarily stable. They are susceptible to the invasion of mutant "parasitic" molecules whose production is catalysed by one link in the cycle but do not continue the cycle by catalysing the next molecule in the hypercycle. Thus, the parasite molecules draw catalytic energy from the cycle. Boerlijst and Hogeweg developed a model which placed cyclic catalytic reactions in the framework of a CA (Boerlijst and Hogeweg 1992). The catalytic interactions of the molecules in their model produced the spiral waves typical of many CA models of excitable media. But the interesting aspect of these waves was that they had the effect of pushing parasitic mutants to the periphery where the parasites did not destroy the integrity of the hypercycles. Boerlijst and Hogeweg's results suggest an important role for spatial structure in the evolution of parasites beyond just the molecular level.

Dispersal in Parasites

Conventional wisdom in parasitology holds that a severe disease is an indication of a recent evolutionary coupling of the host and parasite species. It is assumed that, with time, the parasite will evolve to become more benign (Ewald 1983; Anderson and May 1982). However, Ewald argues that disease severity can be better understood in light of the parasite's method of transmission (Ewald 1983, 1993). When a parasite is transmitted by a highly mobile organism, like a biting arthropod, or is itself highly mobile, it can immobilize or kill its host without jeopardizing its own reproductive success. On the other hand, parasites with less mobility cannot afford to destroy their host because it is likely that the parasite's offspring will also depend on that host and its offspring, or at least depend on the mobility of the host for transmission. Ewald supports this argument by classifying a number of human diseases by severity and showing that the severe diseases tend to have mobile vectors for transmission (Ewald 1983). Similarly, parasites that are transmitted vertically, through the cytoplasm of ova from mother to offspring, rather than horizontally, between unrelated individuals, tend to kill off male hosts, which are evolutionary dead-ends for the parasites, while having relatively benign effects on female hosts (Hurst 1991). Herre has observed a strong correlation between mode of parasitic nematode transmission and the virulence of their effects on their fig wasp hosts (Herre 1993). However, this support is correlational and so requires further experiments to test the association of disease severity with forms of transmission. Through the analysis of mathematical models, Anderson and May conclude that, "...formal studies make it clear that the coevolutionary trajectory followed by any particular host-parasite association will ultimately depend on the way the virulence and the projection of transmission stages of the parasite are linked together: depending on the specifics of this linkage, the coevolutionary course can be toward essentially zero virulence, or to very high virulence, or to some intermediate grade." (Anderson and May 1982)

Tierra

The Tierra model provides a rich platform for experimentally investigating evolutionary dynamics. Tierra simulates evolution in an early RNA world (Szathmari 1989) by drawing an analogy between primitive one-celled organisms and self-replicating Tierran assembly language programs. In Tierra, evolution is based on the reproduction of these programs and the variation that arises in the population of programs due to mutation. The heterogeneous population of programs naturally leads

to a selection dynamic because some of the programs are more efficient self-replicators than others.

A Tierran organism consists of two distinct parts, the code of the assembly language program, and a virtual computer to execute that code. The assembly language program can be thought of, and will often be referred to as, the genome of the Tierran. Ray developed an assembly language specifically for the model because standard assembly language programs tend to crash when subjected to mutation. In other words, small syntactical changes in a standard assembly language tend to have drastic semantic effects. For our purposes, it is only important to note that the language is made up of a small set of instructions (32) that do not take arguments. This means that a genome is a word written in a 32 letter alphabet, and so bears a closer resemblance to a protein, based on an alphabet of 22 amino acids, than to RNA or DNA, based on an alphabet of 4 nucleic acids.

The Tierran virtual computer is a simplified processor. It has two registers for holding addresses of locations in the environment. It also has two registers for manipulating numbers. Furthermore, there is a stack that can be used for temporary storage of information. However, the most important part of the processor is the instruction pointer which holds the address of the next instruction to be executed by the processor.

Those familiar with the Core Wars game (Dewdney 1984, 1985, 1987, 1989) will be familiar with this sort of model. However, Tierra differs significantly from Core Wars on three counts. First, an organism is not allowed to modify, or write into, another organism's genome, except for the genome of its offspring during replication. This is analogous to the protection afforded by a cell wall. Second, the Tierra world contains both mutation and non-deterministic, or erroneous, execution of the program instructions. This leads to variation in the population. Third, selection in this heterogeneous population is implemented by killing off the old and error-prone organisms when the environment fills to 80% capacity.

The Tierra model is powerful because it is simple and yet it captures the essential dynamics of evolution. It is also a relatively open-ended model when compared to traditional genetic algorithms (Goldberg 1989), because there is no fixed length of a Tierran genome and a Tierran program can compute any computable function (Maley 1993). The behaviour of the genome is specified in the genome itself through the interactions of the genes or instructions, rather than being specified through a mapping of gene locations onto aspects of an individual's phenotype.²

Perhaps more striking than these structural innovations is the behaviour of the Tierra model. Ray observed the emergence of a viral style parasite, predator-prey population dynamics, and genetic arms races. One of the most striking and consistent developments in the Tierran evolution is the rise of viral style parasites. These are organisms that have lost a part of their genome and must depend on another organism to supply the missing part. This loss of genetic material results in parasites about half the size of their hosts. The smaller size means that parasites are able to replicate more quickly than their hosts and also make more efficient use of space in the environment. This competitive advantage, paired with the parasite's dependence on the host species, leads to the coupled oscillations between the host and parasite species that are typical of predator and prey species. Finally, an arms race develops between the hosts and parasites, with the hosts evolving immunities and the parasites evolving strategies for circumventing those immunities. It is precisely these dynamics, typical of real ecological communities, that reinforce the analogy between Tierran evolution and biological evolution. For a complete description of the Tierra model see Ray, 1992.

Plateau

In Tierra, the environment is a one dimensional array of integers. Each integer represents a single instruction in Tierran assembly language, along with some information about the owner of that instruction. For our purposes it is only important

²Koza also achieves a similar kind of open-ended evolution with his automatic programming (Koza 1992).

to note that the environment is one dimensional. In this environment, an organism is a contiguous series of integers in the array.

In my expansion of the Tierra model, which I call Plateau, I have added a dimension to the environment. The Plateau environment is a two-dimensional array, or a grid. Furthermore, since all borders "wrap-around" the environment is, in effect, a torus. However, it is not a traditional torus. Organisms in the Plateau model, Platonists, remain one-dimensional, like their Tierran cousins. They are horizontal strings of instruction, with execution generally flowing from left to right. If the environment were a traditional torus, an organism would be limited in length to the circumference of the torus. Any organism longer than the width of the environment, the circumference of the torus, would wrap around and overwrite itself. To avoid this limitation, I shifted the left and right edges by one row. In other words, the rightmost instruction in row n is adjacent to the leftmost instruction in row $n+1$, making the environment resemble a coiled spring with its two ends attached to form a torus.

The expansion of the environment to two dimensions without also allowing the organisms to be two-dimensional³ would not have any effect on the model except for the fact that interactions now occur in two dimensions. When searching for a location, organisms first search their genome. However, if they fail to find the desired location in their own genome, they leave their genome to search through the two-dimensional environment, constrained by a radius of locality.⁴

The other significant addition to Plateau was the introduction of a dispersal parameter for reproduction. In Ray's Tierra, a newborn organism is placed wherever it will fit in the environment. In contrast, reproduction on the Plateau involves the specification of a location for the newborn. If that space is not empty, the current occupants are moved aside to make the required space. This had the unfortunate effect of introducing movement to the model. Movement raises a number of difficult problems⁵ if one wants to maintain the coherence of interactions in the highly connected world of the Plateau. Two forms of dispersal have been investigated on the Plateau. The first form I will call *adjacent dispersal*. *Adjacent dispersal* involves placing the offspring adjacent to the parent, on the left or the right. The second form of dispersal I will call *global dispersal*. In the *global dispersal* condition, offspring are placed in a randomly chosen location in the environment.

The Experiment

Design

The Plateau model was run repeatedly under the *adjacent* or *global dispersal* condition in order to investigate the effects of dispersal on the evolution of the Platonists. Data were collected for 22 trials, 11 under the *adjacent dispersal* condition and 11 under the *global dispersal* condition. Otherwise, the trials only varied by the seed given to the random number generator.

³I plan to make this logical next step soon. However, it involves the significant problems of designing a two-dimensional assembly language that is robust to mutations. The proportion of possible fatal mutations in a program must be reduced in order to increase the speed with which novel self-replicating programs can evolve.

⁴The entire locale is not searched, but rather the organism's pointer is sent in one of the three directions from which it has not come, and allowed to search for a location along that trajectory. Davis spoke of a significant reduction in execution speed if the entire locale is searched (Davis 1992).

⁵Interactions in Tierra and Plateau are based on location pointers in the address registers. If a location pointer is pointing to a gene in another organism, and that organism moves, the pointer must be updated or else it will suddenly point to a different gene. Plateau solves this problem by implementing location pointers that are relative to the organism that own the gene, so that when that organism moves, the location pointer does not have to be updated. However, when that organism dies, the pointer does need to be updated. Since a pointer must be updated every time it moves, this problem can be avoided by keeping the dead organisms in a list of "ghosts" until all of the pointers that were using them as a relative address have moved and been updated. By making addresses relative to the owner organisms, the memory allocation algorithm can be implemented in $O(n)$ time.

A trial consisted of 2×10^7 iterations. An iteration is an execution of a single instruction. To give an indication of the scale of time on the Plateau, it takes between 10^2 and 10^3 iterations for a single organism to reproduce. The carrying capacity of the environment was generally below 10^3 . So a typical trial traced between 10^1 and 10^2 generations of Platonists.

Data Collection

Two forms of data were examined from a Plateau trial. First of all, any genotype was saved that successfully reproduced twice and was unique amongst the successful organisms in the population at the time of the second reproduction. These genotypes were later examined to determine if they were self-sufficient replicators or some form of parasite. Second, every 500,000 iterations, all genotypes were logged that made up more than 1% of the population. The log file consisted of the name of the genotype and the current number of organisms with that genotype. The log files were examined to determine the point at which the relatively successful parasites outnumbered the relatively successful hosts.

Results

In the trials with *adjacent dispersal*, the parasites outnumbered the hosts after 6.91×10^6 iterations on average. In contrast, with *global dispersal*, the parasites outnumbered the hosts after 4.27×10^6 iterations on average. The difference was statistically significant ($p = 0.026$).⁶ Furthermore, the *adjacent dispersal* trials recorded an average of 85.5 successful genotypes, while the *global dispersal* trials recorded only 62.5 successful genotypes on average ($p = 0.013$). When the genotypes were divided into host and parasite genotypes, there was no statistically significant difference between the number of parasitic genotypes recorded under the two conditions ($p = 0.72$), but there was a strongly significant difference between the number of host genotypes recorded: 42.7 for the local dispersal trials, and 21.9 for the *global dispersal* trials ($p = 0.0061$).

Discussion

Parasitism

The *adjacent dispersal* condition is more hostile to the evolution of parasites. When dispersal is adjacent, a successful parent is constantly pushing aside its neighbours and placing offspring in that space. This form of dispersal results in clusters of identical genotypes, with perhaps a few minor differences caused by mutations. Since most parasitism requires some degree of proximity, *adjacent dispersal* will push parasites to the periphery of host communities and thus provide a degree of protection for the hosts near the centre of the patch.⁷ Furthermore, as a successful parasite reproduces, it increases the amount of space between itself and its host and thus may lose its ability to utilize that host.

There was some evidence of lineages losing their resistance to parasitism under the *adjacent dispersal* condition. This might be expected due to their reduced exposure to parasitism and thereby the reduction in selection pressure applied by parasites. However, this loss of resistance was also seen in the *global dispersal* conditions, and the differences were not statistically significant with the sample size.

Adjacent dispersal can be seen as analogous to vertical transmission of the parasite. The offspring of the parasite will most likely be dependent on the parasites

⁶ The data were analyzed with a two-tailed T-test. Differences between populations were considered statistically significant when $p < 0.05$.

⁷ Most parasitism occurs by using the CALL instruction to execute a procedure in another organism's genome. However, the CALL instruction is restricted to searching in a 250 gene radius. Since most hosts were about 50 genes long, it took a buffer of about five hosts (or ten parasites) on either side to protect a host from being parasitized.

host or its offspring. In contrast, *global dispersal* is analogous to horizontal transmission since a parasite's offspring will tend to utilize a different host organism than its parent. We would predict that under the *adjacent dispersal* condition the parasites would evolve towards avirulence since their reproductive success is intimately bound to the reproductive success of their hosts. However, since only one form of parasitism was observed on the Plateau, it was impossible to examine different degrees of virulence in the different conditions.

Genetic Variation

The increase in genetic variation in trials with *adjacent dispersal* can be understood in terms of patchiness in the environment. Since *adjacent dispersal* leads to clusters of organisms with nearly identical genotypes, there is a variety of different local environments in the world. Furthermore, without any dispersal, the offspring of an organism will experience an environment similar to the parent's environment. This allows a lineage to evolve adaptations to its local environment. In this way, *adjacent dispersal* leads to both a patchy environment and a variety of successful lineages adapting to their local environments.

One competing explanation was considered for the observation of increased genetic variance in the *adjacent dispersal* condition. Since the measurement of genetic variation was based on the number of successful organisms recorded, and a successful organism is defined in part by being unique at a specific time, it is possible that the same genotype may have been successful, gone extinct, and then risen again repeatedly. Why would this happen more often in the *adjacent dispersal* condition? An entire genotype might go extinct if it were a colony of parasites depending on a single host. This probably happens more frequently in the *adjacent dispersal* condition than in the *global dispersal* condition. If that host dies off, then the colony may be left stranded without any way to reproduce. This explanation would predict greater "genetic variation" amongst the parasite species, but not amongst the host species. However, the opposite was observed, and so the explanation was rejected.

These results underline the importance of a patchy environment coupled with a demic structure of populations for the development of genetic variation. When there is *global dispersal* on the Plateau, the organisms are fully mixed and so, in some sense the environment is homogeneous. However, the colonies formed under the *adjacent dispersal* condition associate information with locations in the environment. Organisms can then evolve to exploit this information, or, in other words, adapt to their surroundings. It is also important to note that it is the organisms themselves that form the interesting aspects of the environment. There was clear evidence of genetic arms races between host and parasite species. Although, it was also clear that the hosts were at a disadvantage in these arms races since it generally took two mutations in a host genome to develop a resistance to parasitism, while it only took one mutation in a parasite genome to overcome most forms of resistance.

Extensions

It would be interesting to allow the organisms to specify their own mode of dispersal. This could be done by adding a second argument to the memory allocation (MAL) instruction which would be an indication of the distance from the parent that the offspring should be placed. Then we could see if the hosts and parasites diverged, with the parasites evolving to disperse more widely, while the hosts evolve to place their offspring as near as possible in order to develop a protective buffer from parasitism.

By adding dispersal parameters to the model, we have taken a first step to adding a demic structure to the model. However, since all the demes are asexual there is no mixing, with the possible exception of a primitive form of sexuality due to parasites shuffling host's genes (Ray 1992). Without mixing, the clusters of organisms cannot function as gene pools, and migration becomes relatively meaningless. The power of a demic structure could be realized if sexuality were added to the model, as it has recently been added to Tierra (Ray 1992a).

Conclusion

The results from the Plateau model support the hypothesis that the severity of a disease depends on the mode of parasite transmission. *Global dispersal* is generally more amenable to parasitism than local, or in this case, adjacent, dispersal. This support is predicated on the strength of the analogy between the evolutionary dynamics on the Plateau and biological evolution. The addition of spatial structure and the explicit modelling of individuals (Caswell and John, 1992; Huston et al. 1988) can lead to results that are not predicted by more abstract models. In this case, the addition of spatial structure leads first to the prediction that biological parasites will invade and exploit a population more slowly when they disperse their young over short distances than when they disperse their young on a wider scale, and second to the prediction that when biological organisms disperse their young for only short distances more genetic variation will evolve than when organisms disperse their young widely.

References

- Anderson, R.M. and May, R.M. 1982. Coevolution of hosts and parasites. *Parasitology* 85, 411-426.
- Bak, P., Wiesenfeld, K., Tang, C. 1988. Self-organizing criticality. *Physical Review* 38A (1), 364 - 374.
- Boerlijst, M. and Hogeweg, P. 1992. Self-structuring and selection: spiral waves as a substrate for prebiotic evolution. In: Artificial Life II, C. Langton et al. (eds.) Redwood, CA: Addison-Wesley. pp. 255-76.
- Caswell, H. and John, A.M. 1992. From the individual to the population in demographic models. In: Individual-based models and approaches in ecology, D. DeAngelis and L. Gross (eds.) New York: Chapman and Hill. pp. 36-61.
- Davis, Paul. Personal communication. June 16, 1992. He can be contacted at pauld@cs.washington.edu.
- Dewdney, A.K. 1984. In the game called Core War hostile programs engage in a battle of bits. *Scientific American* vol. 250, May.
- Dewdney, A.K. 1985. A Core War bestiary of viruses, worms and other threats to computer memories. *Scientific American* vol. 252, March.
- Dewdney, A.K. 1987. A program called MICE nibbles its way to victory at the first Core War tournament. *Scientific American* vol. 256, January.
- Dewdney, A.K. 1988. The hodgepodge machine makes waves. *Scientific American* vol. 259, August, 104 - 107.
- Dewdney, A.K. 1989. Of worms, viruses and Core War. *Scientific American* vol. 260, March.
- Dyson, F.J. 1985. Origins of Life. Cambridge: Cambridge University Press.
- Ewald, P.W. 1983. Host-parasite relations, vectors, and the evolution of disease severity. *Ann. Rev. Ecol. Sys.* 14, 465-85.
- Ewald, P.W. 1993. The Evolution of Virulence. *Scientific American* vol. 268, April, 56 - 62.
- Goldberg, D.E. 1989. Genetic Algorithms in Search, Optimization, and Machine Learning. Reading, MA: Addison-Wesley.
- Herre, E.A. 1993. Population structure and the evolution of virulence in nematode parasites of fig wasps. *Science* 259, March 5, 1442 - 1445.
- Hurst, L.D. 1991. The incidence and evolution of cytoplasmic male killers. *Proc. R. Soc. Lond. B* 244, 91 - 99.
- Huston, M., DeAngelis, D., and Post, W. 1988. New computer models unify ecological theory. *Bioscience* 38(10), 682-691.

- Koza, J.R. 1992. Genetic evolution and co-evolution of computer programs. In: Artificial Life II, C. Langton et al. (eds.) Redwood, CA: Addison-Wesley. pp. 603-30.
- Langton, C. G. 1986. Studying artificial life with cellular automata. *Physica D* 22, 120 - 149.
- Maley, C.C. 1993. The Computational Completeness of Ray's Tierra Assembly Language. In: Artificial Life III. C. Langton et al. (eds.) Redwood, CA: Addison-Wesley. In press.
- Ray, T.S. 1992. An approach to the synthesis of life. In: Artificial Life II, C. Langton et al. (eds.) Redwood, CA: Addison-Wesley. pp. 371-408.
- Ray, T.S. 1992a. Haploid sex. *Tierra Digest* vol. 1, num. 9. July 8. An email publication. For information or subscription email: tierra-digest@life.slhs.udel.edu.
- Szathmari, E. 1989. The emergence, maintenance, and transitions of the earliest evolutionary units. *Oxford Surveys in Evolutionary Biology* 6, 169 - 205.