Rapid Host-Parasite Coevolution Drives the Production and Maintenance of Diversity in Digital Organisms

Luis Zaman
Department of Computer Science and Engineering
Michigan State University
East Lansing, MI, USA
zamanlui@msu.edu

Suhas Devangam
Microbiology and Molecular Genetics
Michigan State University
East Lansing, MI, USA
devangam@msu.edu

Charles Ofria
Department of Computer Science and Engineering
Michigan State University
East Lansing, MI, USA
ofria@msu.edu

ABSTRACT

Accumulating evidence suggests evolution and ecology can happen on similar time scales. Coevolution between hosts and parasites is a practical example of interacting ecological and evolutionary dynamics. Antagonistic interactions theoretically and experimentally increase host diversity, but the contribution of novel variation to diversity is not well understood. In laboratory or natural settings it is infeasible to prohibit novel mutations in communities while still allowing frequencies of extant organisms to change. We turn to digital organisms to investigate the effects of rapid evolution on host-parasite community diversity in the presence and absence of novel variation. We remove the source of variation in coevolved digital host-parasite communities and allow them to reach an equilibrium. We find that coevolved host-parasite communities are surprisingly stable in the absence of new variation. However, the communities at equilibrium are less diverse than those that continued to experience mutations. In either case, hosts coevolving with parasites are significantly more diverse than hosts evolving alone. Harnessing an advantage of in silico evolution, we show that novel variation increases host diversity in communities with parasites further than the trivial increase expected from new mutations.

Categories and Subject Descriptors

F.1.1 [Computation by Abstract Devices]: Models of Computation—Self-modifying machines

General Terms

Experimentation

Keywords

Evolution, Coevolution, Parasites, Antagonism, Diversity, Digital Evolution, Artificial Life, Rapid Evolution

1. INTRODUCTION

Theodosius Dobzhansky’s now famous words, “Nothing makes sense in biology except in the light of evolution...” are especially true for the problem of biodiversity [7]. Evolution is the process that shaped all of life, extant and extinct. However, biodiversity is typically thought of in an ecological framework, interested in how a static set of species with static interactions can be stably maintained. Evolution is instead reserved for grand scales, defined by Slobodkin as happening on the order of half a million years [25]. Contrary to this view, many studies have revealed substantial evolution occurring over very short time scales; this concept of rapid evolution is reviewed in [12, 23, 10, 5]. Sometimes evolution occurs on such short time scales that it has significant ecological effects on communities [28, 17]. This new view of eco-evolutionary dynamics means the problem of biodiversity must include the feedbacks between ecology and evolution in order to understand how rapid evolution influences the maintenance of diversity in communities [26].

Parasitic species are paramount to the biodiversity found in nature with nearly half of all known species classified as parasites [20, 21]. They are not only incredibly diverse and successful, but parasites can theoretically and experimentally increase host diversity [8, 4, 2, 9, 24]. Similarly, the escape from natural parasites that would otherwise limit growth is a leading hypothesis for why some introduced species become invasive [27]. For these and many other reasons, parasites can be used counterintuitively as indicators of ecosystem health, where communities with a high diversity of hosts often have many parasites [13]. Another counterintuitive effect of host-parasite coevolution is the increase in protection against emerging diseases it provides. Host diversity is a key factor in mitigating novel pathogens and their potential for rapid evolution, but a major source of genetic diversity results from coevolution between hosts and parasites. Thus, in order to prevent emerging disease disasters, we must learn how to protect and foster the coevolution of hosts and their parasites [1]. Antagonistic coevolution also offers the perfect situation to study rapid evolution effects, since ecological feedbacks are inherent in their evolutionary dynamics [3]. We are generally interested in how rapid coevolution creates and maintains diversity in host-parasite communities, and specifically interested in how novel variation affects the resulting diversity. Even when studying rapid evolution, several generations can take years to observe and performing detailed experiments in a natural setting is infeasible. Thus,
many examples of rapid evolution come from experimental evolution of microbes [23]. However, even these microbial systems have drawbacks when studying rapid evolution, such as difficulties in assessing the entire population of interesting traits, and the inability to control for random processes like mutation. In order to investigate the role of novel variation in host diversity, we turn to artificial communities.

To study the coevolution of host-parasite communities in silico, we use the Avida research platform [19, 18]. We implement parasitic organisms and a mechanism for them to infect hosts based on genetically encoded phenotypes. We compare independent populations of digital organisms in the presence and absence of these parasites, as well in the presence and absence of novel variation taking the form of random mutations. We find that hosts coevolving with parasites are more diverse than hosts evolving alone. This result holds both in the presence and absence of novel variation. However, new variation increases diversity in host-parasite communities more than it does for hosts in the absence of parasites, suggesting the importance of novel variation in maintaining stably diverse communities.

2. MATERIALS AND METHODS

2.1 Avida

For all experiments, we use Avida 2.13.0 r4173. Avida is a digital life research platform that maintains a population of self-replicating computer programs (“digital organisms”), which compete for resources. Digital organisms exist in a well-mixed environment, interacting randomly with any other organism in their world. Genomes consist of a circular list of instructions from a Turing complete programming language, executed on virtual hardware. Each instruction directs the organism to perform a simple operation such as arithmetic, flow control, or environmental interaction.

During replication, an organism loops through its genome copying each instruction sequentially until reaching its end. It then executes a divide instruction, separating off its offspring and placing it into a random cell in the world, replacing any previous occupant. Instead of always producing a perfect replica of the parent genome, the copying process is noisy and introduces errors. These mutations and can be insertions, deletions, or substitutions.

Organisms replicate by using their virtual CPU to execute an appropriate series of instructions. In these experiments, there is a single type of resource that must be metabolized for successful replication. To metabolize a portion of the available resource, organisms perform logical tasks on environmental inputs. The default Avida environment contains nine logic tasks. In these experiments, all nine default tasks are available, but they metabolize the same resource assuming there is a sufficient quantity available (See section 2.3).

Organisms in Avida possess one of several virtual hardware types, varying in instruction set and architecture. In the hardware type we use for this study, organisms have: four stacks to store and manipulate numerical values, a set of genome memory spaces in which organisms execute and copy instructions, and a set of heads that point to positions in each memory space. Organisms identify a specific stack, memory space, or head with a label consisting of no-operation instructions (“nops”).

Since Avida has heritable variation, and environmentally driven selection, evolutionary dynamics are a natural byproduct of the system [6]. For this reason, Avida has been used successfully to understand ecological and evolutionary dynamics, as well as to perform more applied research in distributed systems and software engineering. For a detailed introduction to Avida, see [18, 19].

2.2 Parasites

Parasitic digital organisms are self-replicators that operate inside hosts, relying on them to provide energy in the
form of CPU cycles. Instead of executing a \texttt{divide} instruction to finish replicating, a parasite must inject it’s offspring into a host. When the \texttt{inject} instruction is executed, the parasite offspring attempts to infect the organism in a randomly chosen location. If successful, the new parasite is treated like a thread in the host organism, consuming CPU cycles and thus reducing it’s host’s fitness. Infection is successful if any of the logical tasks performed by the parasite match any of the tasks the host is performing. Infection will fail if there is no overlap in tasks, if the chosen location is empty, or if the organism is already infected. The probability of a parasite stealing a CPU cycle from its host is configurable, and we will refer to it as “virulence”. When virulence is set to one, parasites steal all CPU cycles from their hosts, killing them and using them for energy like predators. Indeed, when observing the ecological dynamics of parasites with maximal virulence, we find classic Lotka-Volterra dynamics (Figure 1) [14]. When virulence is set to 0.5, parasites and hosts split CPU cycles evenly and there is a stable equilibrium of host and parasite frequencies.

Hosts and parasites in Avida are similar to \textit{E. coli} and lambda phages. These phages must attach to receptors on the surface of bacteria in order to infect, and the bacteria must have receptors in order to consume resources from the environment. However, consuming resources leaves them susceptible to phages. The bacteria evolve resistance by changing their surface receptors, but lambda phages can counter resistance by evolving their tail fibers to attach to these new proteins [16]. Similarly, in Avida, hosts must perform logic tasks to consume resources and thus replicate, but this action leaves them susceptible to infection. Resistance can evolve by changing the logic task(s) used to consume resources, but the parasites can counter adapt by evolving the ability to perform the new task. Figure 2 depicts the mechanics of infection in Avida.

\section{2.3 Configuration}

All experimental runs are in a well-mixed environment, where host and parasite offspring are randomly placed in the world. Each run starts with a 320-instruction-long host organism capable of performing only the \texttt{NOT} task and self-reproduction. In runs with parasites, after 3,000 updates 400 cells in the world are exposed to 80-instruction-long ancestral parasites capable of performing only \texttt{NOT} and self-reproduction. The parasites in these experiments have a virulence of 0.80 unless otherwise noted. In order to become resistant, hosts must loose their ancestral task (so that the parasites cannot infect them) while also evolving a novel task (so that they can continue to collect the resources required for replication). Similarly, in order for parasites to infect hosts that evolve resistance, they must also evolve the novel task. These ancestral hosts and parasites are capable of performing only the most basic task in the environment and self-replicating, the rest of their genomes are padded with no-operation instructions. Each run is allowed to execute for 200,000 updates, where one update is the amount of CPU time needed for each organism in the population to execute an average of 30 genomic instructions.

We disallow multiple infection by setting the maximum number of threads an organism can have to two. Organisms are not given access to instructions manipulating or creating their own threads, preventing host organisms from becoming resistant by simply creating additional threads. We also disallow vertical transmission, the direct inheritance of parental parasites, by clearing infections on successful host division. Only allowing horizontal transmission prohibits the association between parasite reproduction and host reproduction, since we consider cell division the birth of two daughter-cells.

Novel variation can take the form of point, insertion, and deletion mutations. These mutations are applied per site, and are split 10:1 point mutations to insertions/deletions. The individual rates are set so that, on average, hosts will have a single mutation every four offspring, and parasites will have a single mutation every two offspring.

There is a single resource that must be consumed by hosts for successful replication. This resource is kept at a low level to prevent the world from filling: there are 14,400 potential locations for organisms and approximately 10,000 are filled in the absence of parasites. For a host to successfully consume this resource, there must be a sufficient quantity in the environment. For these experiments we require two units to be available, and one unit is consumed. If a host does not successfully consume resources before executing the \texttt{divide} instruction, replication will fail and the organism will begin execution again without producing any offspring. Eventually, organisms will die of old age if they do not successfully divide (since asexual reproduction produces two daughter cells, age is reset on successful division) and the maximum age in these experiments is $30 \times$ genome length. Cell death may also occur as a result of offspring being randomly placed in the world. In this case, the occupant is overwritten by the newly divided cell.

\section{2.4 Measuring Diversity}

We measure diversity as the Shannon diversity index ($H$) of binary phenotypes. That is, we look only at whether each of the nine tasks is performed or not without accounting for expression level and consider each unique binary string a

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure2.png}
\caption{A diagram of the traits governing host-parasite interactions. The green square depicts the single resource type in the environment that must be consumed for successful host replication. Hosts can use any of the nine default logic tasks, indicated in blue, to consume part of this resource if it is available. Parasites, depicted as red triangles, target the mechanism hosts use to consume resources - the logic tasks.}
\end{figure}
Figure 3: Eco-evolutionary dynamics of host-parasite interactions. Subfigures (a) and (b) are respectively a typical frequency plot and phase plane of host-parasite interactions in the absence of evolution - the ecological dynamics. Subfigures (c) and (d) are the otherwise identical, except that mutations are allowed, and thus evolution can occur. Both subfigures are of approximately 100 generations, representing an ecological time scale. These figures demonstrate the disruption of typical ecological dynamics (compare (a) to (c) and (b) to (d)) by evolution on ecologically relevant time scales.

2.5 Measuring the Effect of Parasites on Host Diversity

To compare the overall effect of parasites on host diversity, we ran 50 replicate populations where we introduced parasites and 50 replicates where we did not. We then measured the Shannon diversity index of the final set of hosts as described in Section 2.4. The difference in host diversity between these two treatments is the effect coevolution with parasites had in these communities.

We also measured how parasites influenced diversity in ecological communities, where no new variation was being introduced, by running 50 replicate populations that contained parasites, and 50 replicates that did not. All runs (co)evolved for 100,000 updates, then we disallowed mutations and the communities continued for an additional 100,000 updates to settle into an equilibrium. We then compared the Shannon diversity index of hosts from the resulting communities in these two treatments to quantify the parasites’ ecological contribution to diversity.

Note that both equilibrium and ecological are misnomers, though in this case they are the most proximate terms. Parasites are frequently in a state of non-equilibrium, even in ecological contexts, and populations are technically evolving if there are changes in mean phenotype over time – which can happen in a community context even in the absence of new variation.

2.6 Measuring the Effect of Novel Variation on Host Diversity

To control for novel variation in host-parasite communities, we harnessed the repeatability of independent evolutionary runs in Avida. Essentially, we asked what would have happened if we went back in time and allowed the communities to continue mutating by having ensured identical coevolutionary histories. We determined the effect of novel variation by measuring the difference in host diversity between paired runs where novel variation was continued versus when it was stopped.

Unfortunately, there are confounding effects when measuring the contribution of novel variation this way, since new variation is obviously a source of diversity. To correct for this effect, we compared the increase in diversity of communities with parasites to that of communities without parasites (also pairing runs of only hosts as described above). If diversity increases further with parasites than hosts evolving alone, we know that the effect of novel variation in communities with parasites is not due to the trivial rise in diversity new variation brings about. Additionally, the new variation in communities without parasites can still have some influence in diversity above the trivial effect, but these analyses are conservative with respect to the actual contribution of novel variation.
Figure 4: Host diversity in runs that evolved without parasites compared against host diversity in runs that coevolved with parasites. Subfigure (a) depicts host diversity when all 200,000 updates had mutations, and subfigure (b) depicts host diversity when mutations were stopped at 100,000 updates. Thus, subfigure (b) shows the ecological effects parasites have on host diversity.

3. RESULTS

The signature of rapid evolution is the disruption of typical ecological dynamics. Figure 3(a) and 3(b) show the frequency and phase plane plots for a typical host-parasite community without evolution, describing the pure ecological dynamics for this system over approximately 100 generations. Figure 3(c) and 3(d) depict the dynamics in the presence of novel variation and thus evolution. If evolution occurs on different time scales than ecology, we would not expect differences over such few generations. However, comparing 3(a) to 3(c) and 3(b) to 3(d) shows a pronounced difference in community dynamics, suggesting that coevolution occurs on ecological time scales in this system.

3.1 Effect of Parasites

When considering any replicate runs with parasites, we removed communities that lost parasites in either the presence or absence of mutations from analysis (18 runs in total were removed out of 100). In other words, to be considered in the analysis, parasites had to persist in both the 50 original runs, as well as the 50 replays where mutations were continued. Only one of the paired-runs withheld from analysis came from a community that maintained parasites in the absence of novel variation, but lost them when the runs were replayed with mutations. A single community lost parasites after novel variation was removed, but they were also lost when mutations were continued. Thus, the loss of parasites in this case was not due to instability after stopping novel variation. Seven other communities lost parasites prior to losing novel variation. Thus, these seven runs, as well as their seven paired replayed runs were removed from analysis. These data suggest coevolved communities are able to maintain parasites robustly in the absence of novel variation.

Figure 4(a) depicts the Shannon diversity distributions of hosts in communities with and without parasites. Communities with parasites have significantly more diversity (Mann-Whitney U = 1996, p < 0.001). The presence of parasites results in an increase of host Shannon diversity by 1.784 with a 95% confidence interval of [1.506, 2.063].

To measure the ecological effects parasites have on host diversity, we remove the possibility for novel mutations and measure the resulting communities’ host diversity. Figure 4(b) depicts the Shannon diversity distributions of communities with and without parasites after 100,000 updates without mutations. Again, communities with parasites have significantly higher diversity than those without parasites (Mann-Whitney U = 1816, p < 0.001). In these ecological communities, parasites increase the Shannon diversity by 1.15 with a 95% confidence interval of [0.933, 1.434].

3.2 Effect of Novel Variation

The pairwise subtraction of host diversity between runs where novel variation was stopped at 100,000 updates from the runs where mutations continued produces the distribution of increases in Shannon diversity that would have occurred had the ecological community not lost their sources of novel variation. We pair these runs since the communities at 100,000 updates are identical, and thus the only difference is whether or not mutations continued. Figure 5 depicts the distribution of increases in host diversity due to novel variation in communities with and without parasites.

The statistical difference between increases in diversity with and without parasites is a measure of how novel variation affects host diversity in the presence of parasites above and beyond the trivial effects of new mutations. There is a significant difference between these two distributions, where communities with parasites have a 0.652 increase in host diversity with a 95% confidence interval of [0.321, 0.973] (Mann-Whitney U = 1508, p = 0.00012).

4. DISCUSSION

Evidence of evolution happening rapidly enough to co-occur with and influence ecological dynamics is now widely documented [5]. Host-parasite coevolution is intimately connected to ecological dynamics, and thus a likely candidate for these eco-evolutionary feedbacks. Despite the large amount of evidence showing ecological and evolutionary dynamics
Figure 6: Frequencies of phenotypic traits in hosts and parasites for a sample coevolutionary community where mutations stop at 100,000 updates (grey line). Subfigure (a) depicts the relative frequencies of traits hosts use to consume resources in the community through time, while (b) depicts the relative trait frequencies parasites use to infect hosts. Parasites tracking host phenotypes is apparent in the similarity between the phenotype heat maps. Subfigure (c) depicts the frequencies of hosts and parasites through time.

interacting, little is known about how important this feedback is for maintaining communities. Novel variation is only one aspect of rapid evolution, but it is an important one nearly impossible to test in laboratory or natural settings.

We have presented an *in silico* instance of coevolution and demonstrated ecological and evolutionary dynamics interacting on similar time scales. If ecological processes were happening much faster than evolution, we would not expect to see the rapid disruption of typical community behavior when we introduce novel genetic variation to the community. However, as Figure 3 depicts, novel variation did indeed disrupt the ecological dynamics. To understand the effect that this eco-evolutionary feedback had on host diversity in a community context, we first quantified the effects of parasites.

Regardless of whether there was a source of novel variation, parasite presence significantly increased host diversity in this system, consistent with empirical and theoretical results [8, 4, 2, 9]. There are multiple, non-exclusive mechanisms by which parasites can increase host diversity (see discussion in [4]). Parasites may target the most frequent host phenotype, and the hosts would then evolve resistance against these parasites. However, as the new resistant host increases in frequency, parasites would experience selection to target it. Thus, parasite-imposed negative frequency-dependent selection can maintain diversity by keeping the frequency of any one particular host genotype (or phenotype) at bay. Alternatively, hosts may have trade-offs between resistance and competitive ability. When parasites are common, the cost of being resistant may be worth the competitive disadvantage, while at low parasite densities being competitive is more advantageous than being resistant to infection. Again, negative frequency-dependent selection would maintain host diversity. We plan to disentangle these mechanisms maintaining host diversity in the future.

The stability of these coevolved host-parasite communities is surprising. Of all 50 replicate communities that experience the loss of new mutations, only one completely excludes parasites after losing novel variation (a handful eliminated parasites prior to the loss of new variation). In other words, 41 out of 50 host-parasite communities were able to persist both in the presence and absence of continuing novel variation. Figure 6 depicts a single run where all mutations stop at 100,000 updates. Moving from left to right on the figure represents going forward in time, and the heat maps show the relative frequencies of each task being performed by hosts (a) and by parasites (b). In the first 100,000 updates, there is rapid change in host and parasite phenotypes, but as time goes on the variation saturates. This saturation is also evident in 6(c), which depicts host and parasite frequencies. Interestingly, after mutations stop, host and parasite frequencies appear to reach an *equilibrium*, but phenotypes are still changing through time. Additionally, it is clear that neither hosts nor parasites collapse into just one or two phenotypes, rather community diversity persists.

Having shown that parasites increased host diversity both in the presence and absence of novel variation, we aimed to distinguish the effects that mutations had on diversity in host-parasite communities. Novel variation trivially increased diversity, since it often produced new phenotypes.
The relationship between host and parasite diversity and the linear regression for the two treatments of novel variation shows the plot of host diversity versus parasite diversity, interaction of new variation on community dynamics. Figure 7 shows the plot of host diversity versus parasite diversity, and the linear regression for the two treatments of novel variation. The relationship between host and parasite diversity in runs without novel variation for 100,000 updates is strong (Adjusted $R^2 = 0.70$). On the other hand, when novel variation was continued throughout the run, the relationship is much weaker (Adjusted $R^2 = 0.18$). The amount of variation unexplained by the relationship of host and parasite diversity in runs that continued to experience novel variation suggests that additional community or evolutionary dynamics are influencing host and parasite diversity. The large amount of explanatory power the relationship has in the absence of novel variation adds support to this view. Understanding the mechanisms acting on this variation to produce non-trivial increases in host diversity will shed light on the important eco-evolutionary interactions shaping community dynamics [23].

5. ACKNOWLEDGEMENTS

This work benefited tremendously from interactions with David M. Bryson, Ben Kerr, Ian Dworkin, Justin Meyer, Beth Miller, and other members of the BEACON Center for the Study of Evolution in Action. Support for this work was provided in part by National Science Foundation grants CCF-0643952 and CNS-0751155. Luis Zaman was supported by an AT&T Labs Fellowship. Suhas Devangam was supported by the Professorial Assistantship Program through Michigan State University’s Honors College.

This material is based in part upon work supported by the National Science Foundation under Cooperative Agreement No. DBI-0939454. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

6. REFERENCES


