

Digital Evolution in Time-Dependent Fitness Landscapes

Ye Li

California Institute of
Technology
Pasadena, CA 91125
yel@its.caltech.edu

Claus O. Wilke*

Digital Life Laboratory 136-93
California Institute of
Technology
Pasadena, CA 91125
wilke@caltech.edu

Abstract We study the response of populations of digital organisms that adapt to a time-varying (periodic) fitness landscape of two oscillating peaks. We corroborate in general predictions from quasi-species theory in dynamic landscapes, such as adaptation to the average fitness landscape at small periods (high frequency) and quasistatic adaptation at large periods (low frequency). We also observe adaptive phase shifts (time lags between a change in the fitness landscape and an adaptive change in the population) that indicate a low-pass filter effect, in agreement with existing theory. Finally, we witness long-term adaptation to fluctuating environments not anticipated in previous theoretical work.

Keywords

Digital organisms, fluctuating environment, quasi species

1 Introduction

Time-dependent (or fluctuating) environments have attracted a considerable amount of interest in biology, in particular in the ecological literature. Within ecology, environmental variation can affect biological processes and population structures [11], and it can be important for the evolution of ecosystems. For example, it is plausible that constant or slowly changing environments favor the evolution of specialists, while fast changes in the landscape foster the emergence of generalists [22]: In a constant environment, generalists are expected to lose unselected functions, as a result of mutation accumulation, antagonistic pleiotropy (adaptations to the peculiarities of the environment have a negative effect on unselected functions), or both. Therefore, over time specialists will prevail. In a rapidly changing environment, specialists will fare better than generalists only for short periods of time, and fare worse at all other times. Therefore, in the long run generalists will prevail. Within the field of population genetics, changing environments have been shown to favor the evolution of recombination [1] and affect mate choice in favor of heterozygotes [2].

On the level of the individual sequence (or groups of related sequences), fluctuating environments have been shown to affect the evolution of mutation rates [10, 7, 15], to a level where the population can just about follow the changes in the environment [9].

There is general interest in the effect of fluctuating environments on populations of macromolecules evolving at high mutation rates [14, 16, 29], because these models of evolution are tractable either analytically or computationally, and make definite predictions. However, it is difficult to test these predictions experimentally, because adaptation is a slow process, and adaptation to temporally varying environments even more so. While the rise of mutator alleles (predicted by theory [13, 7, 21]) has been ob-

* Corresponding author.

served experimentally [20], few other predictions are amenable to experimental testing with standard experimental organisms.

In this contribution, we set out to test several predictions of Eigen's theory of macromolecular evolution [5, 6] extended to dynamic (meaning time-dependent) fitness landscapes [29], in particular fitness landscapes that vary periodically. As our experimental organism of choice we use digital organisms of the Avida variety (see [18] in this issue for an overview and user's guide, and [26] for a review of recent work using digital organisms).

Avidians have been shown to display a kind of adaptive dynamics ("survival of the flattest" [30]) that is one of the hallmarks of a quasi species. Quasi species are the closely related groups of organisms, localized around one or a few high-fitness types, that are predicted by Eigen's theory to dominate populations at high mutation rates. In Avida, mutation rates can be set by the experimenter, and we chose here two different rates ($\mu = 0.1$ and $\mu = 1.0$) that can be characterized as fairly high and high, and that are comparable to viral mutation rates. Thus, we believe that the experimental conditions are such that Eigen's theory applies.

There are a number of unique predictions of Eigen's theory in dynamic fitness landscapes [14, 16, 29] that are amenable to experimental tests. The general response to a changing peak depends on the frequency and the type of that change. If the change is slow, the population will adapt to the changed height of the peak (quasistatic adaptation), or even move away from its peak to another one. If the change is fast, we expect the population to adapt to the average of the landscape over a period [29, 27]. These predictions can be summarized with the concept of *low-pass filtering*: Evolution acts as a low-pass filter on the population, such that the response of a population to fitness fluctuations at high frequencies is depressed in comparison with the response to the same fluctuations at low frequency. As a consequence of low-pass filtering, we expect a time lag between the movement of the peak and the response of the population, termed the *phase shift*. The low-pass filter effect has previously been seen in one-dimensional models of population dynamics [7, 3, 12] and in quasi-species evolution [16].

Most theoretical investigations of adaptation to time-varying environments discuss the adaptation either to a single peak [28, 16, 27], or to a pair of peaks [29] whose height is changing periodically, or to a peak whose position in genetic space is changing periodically or stochastically [14, 19, 15]. In Avida, landscapes can be carved by specifying the amount of energy (in the form of bonus CPU time) an organism harvests for carrying out logical/mathematical operations. (These operations are called *tasks* [18].) We can create a single oscillating peak in Avida by changing the reward for one particular task only (while keeping all other rewards fixed). Likewise, we can implement stochastically jumping peaks by increasing the reward from zero to a finite value for a stochastically chosen task from a set, while keeping all others at zero. Here, we study an oscillating landscape that is a cross between oscillating and jumping peaks, namely the case of two peaks whose height changes periodically, from full height to flat (absent). As a consequence, this situation can also be viewed as that of a peak periodically jumping in fitness space. We investigate two different pairs of peaks that differ in how difficult it is to evolve (or re-evolve) each associated task. This difficulty scales approximately with the minimum amount of code necessary to complete the task. In Figure 1, the two pairs are depicted schematically, where **NAnd/Not** represents the easier of the task pairs, while **And/OrN** is harder. (Tasks **Echo** and **Add** are always rewarded with constant bonuses. They are stepping stones for the more complex tasks **NAnd**, **Not**, **And**, and **OrN**, and facilitate the evolution of the latter tasks.) More detail on the nature of these logical tasks and how they are implemented in Avida can be found in [18] in this issue.

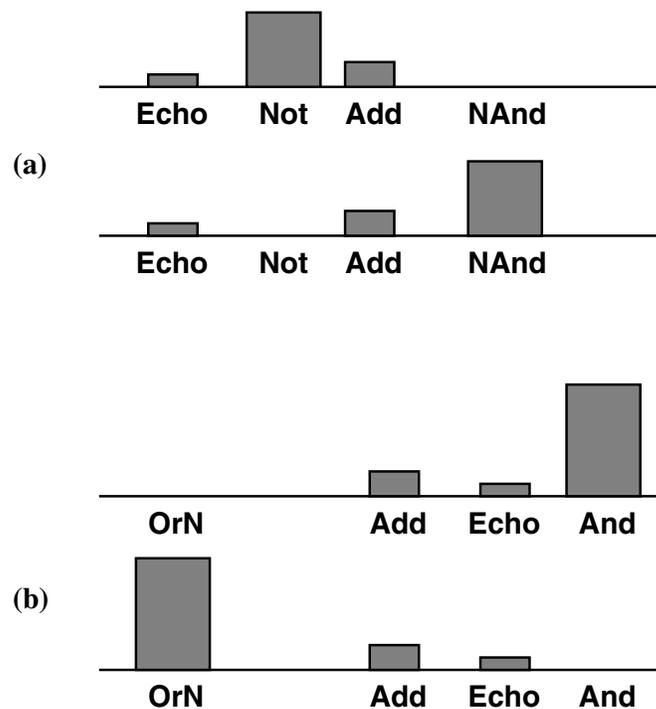


Figure 1. Conceptual view of the landscape with oscillating peaks. (a) NAnd/Not pair of peaks: first half period (upper part), second half period (lower part). (b) And/OrN pair. The location of the peaks for Echo and Add is arbitrary; the heights of the peak qualitatively reflect the awards given. The difference in distance between the Not/Nand peaks and And/OrN peaks is qualitative only and reflects past experience with these tasks.

2 Materials and Methods

All experiments were conducted using Avida version 1.99, available from <http://www.sourceforge.net/projects/avida>. All parameters in the `genesis` file (see [18]) except those mentioned below were kept at their default settings. The size of the world was set to 50×50 (carrying capacity of 2500 organisms). We forced organisms to remain at a constant length of 100 instructions, by setting `CHILD_SIZE_RANGE` to 1 in the `genesis` file. This setting ensures that all offspring are the same length as their parents, and guarantees a constant genomic mutation rate and fairly consistent generation time. We seeded the ancestor runs (see below) with an organism consisting of a copy loop and enough `NOP` instructions to fill up the rest of the genome. Mutations in these experiments occurred only during the copy process.

2.1 Ancestor Runs

In order to limit the effect of history in this study, we evolved different ancestral organisms to seed the later experiments with time-dependent fitness landscape. For each pair of tasks and each mutation rate, we evolved five independent ancestral organisms (for a total of 20). In each replica, we chose the highest fitness candidate to seed the runs with time-dependent fitness landscape.

We chose to set up the reactions using limited resources (instead of the global resourceless fitness landscape), so that rewarded tasks could easily be changed in the events file (as opposed to changing the environment file). A sample environment file for the **Not/NAnd** set of tasks is shown in Figure 2. The **And/OrN** ancestors were

```

RESOURCE resECHO:initial=1000000000000
RESOURCE resADD:initial=1000000000000
RESOURCE resNOT:initial=1000000000000
RESOURCE resNAND:initial=0
REACTION ECHO  resource=resECHO:value=0.2  requisite:max_count=4
REACTION ADD   resource=resADD:value=0.5   requisite:max_count=2
REACTION NOT   resource=resNOT:value=3.0   requisite:max_count=1
REACTION NAND  resource=resNAND:value=3.0  requisite:max_count=1

```

Figure 2. The environment file for the Not/NAnd set of tasks.

created with an identical setup except for the tasks rewarded. We rewarded **Echo** and **Add** in order to facilitate the development of the more complicated tasks. These tasks were kept constant throughout the runs and were performed very consistently by all the organisms.

2.2 Periodic Landscape Runs

Periodic fitness landscapes were implemented via the *events* file. A sample events file for the **Not/NAnd** task pair is shown in Figure 3 for period $T = 2$.

The events file reflects the following protocol. The population is first allowed to equilibrate for 1000 generations. Starting with generation 1001, we switch between the two peaks at constant intervals. In the case depicted in Figure 3, this consists of

```

u 0 inject_all START.CREATURE

# Print standard data files
g 1:1:end print_average_data
g 1:1:end print_dominant_data
g 1:1:end print_tasks_data

# Setup oscillating environment
g 1001:2:end set_resource  resNOT 0
g 1001:2:end set_resource  resNAND 1000000000000
g 1002:2:end set_resource  resNOT 1000000000000
g 1002:2:end set_resource  resNAND 0

# Setup exit time
g 1200      exit

```

Figure 3. The events file for the Not/NAnd set of tasks and a period of 2.

switching **NAnd** for **Not** as the major available resource in the first half of the period, and then switching back in the second half. We varied the period from 2 to 2000 generations ($T = \{2, 4, 10, 20, 40, 100, 200, 400, 1000, 2000\}$). For short periods, we let the run continue for 100 periods. For periods of length $T = 100$ and above, we let the run continue for only 20 periods. We carried out one run for each period length. All experiments were started with populations grown from genotypes adapted to the fitness peak that is present in the first half of the period: **Not** for the easy pair, and **And** for the hard pair.

3 Results

The general dynamics that we observed can be summarized as follows. If the landscape changes very rapidly, the population experiences an average landscape that consists of both peaks. If the landscape changes very slowly, on the other hand, then the population adapts to the currently rewarded task, and loses the unrewarded one (quasistatic adaptation). For intermediate period lengths, the population can maintain both tasks, but during the time interval in which a task is not rewarded, its gene frequency in the population declines. (Here, we refer to the part of an organism's genome responsible for the completion of a particular task as a *gene*, and to the number of organisms that can perform a particular task as the *gene frequency* of that task.)

Figure 4 shows a typical plot of gene frequencies. In the upper panel, the period is small ($T = 2$). We see that the **NAnd** task reaches capacity after approximately 50 generations (=25 periods). (Note that the organism with which we seeded the population was evolved in an environment that rewarded only **Not**, so that it could perform **Not** but not **NAnd**.) Once **NAnd** has reached capacity, however, it is firmly established, because the period during which the gene is not under selection is too brief to lead to its decay. For intermediate period lengths (middle panel of Figure 4, $T = 100$), **NAnd** reaches capacity faster (after approximately 25 generations). Since the period length is now 100 generations, **NAnd** is constantly rewarded for the first 50 generations, so that the selective pressure to perform **NAnd** is stronger than in the case of $T = 2$. Gene frequencies drop somewhat during the unselected periods, because mutations accumulate in the unselected gene. As the period of changes increases further, genes not selected can go extinct, so that they have to be newly acquired in the following period. This effect is more pronounced at high mutation rates (Figure 5), as unselected genes degrade more rapidly then.

The dynamics for each task pair and each mutation rate can be summarized by averaging the gene frequency over all the periods, and plotting them as a function of the phase $\Phi = t(\bmod T)/T$. (Note that with this definition, Φ runs from 0 to 1, and not to 2π .) Figure 6a shows the average **Not** gene frequency for small, medium, and large periods and at mutation rate $\mu = 0.1$. We can see that only for the largest period tested do gene frequencies drop to about half at the end of the unselected period, only to rebound very quickly if selection is reinstated. The rate at which the unselected task decays is much higher in Figure 6b, which shows the same experiment at a tenfold higher mutation rate.

The dynamics for the pair **And/OrN** is qualitatively similar to that of the **Not/NAnd** pair. However, genes are lost quicker and rediscovered slower, a result of the increased difficulty to evolve these tasks. As a consequence, some effects are more pronounced in the data for this task pair than for the easier pair. For example, we notice in Figure 6c that the variance of the gene frequency is unusually large (but constant) for the intermediate period $T = 100$. We can trace this effect back to accidents in evolution: If the task is lost in some periods and *not* reacquired in the next (as sometimes happens), the distribution of gene frequencies at a given phase Φ over all periods is not

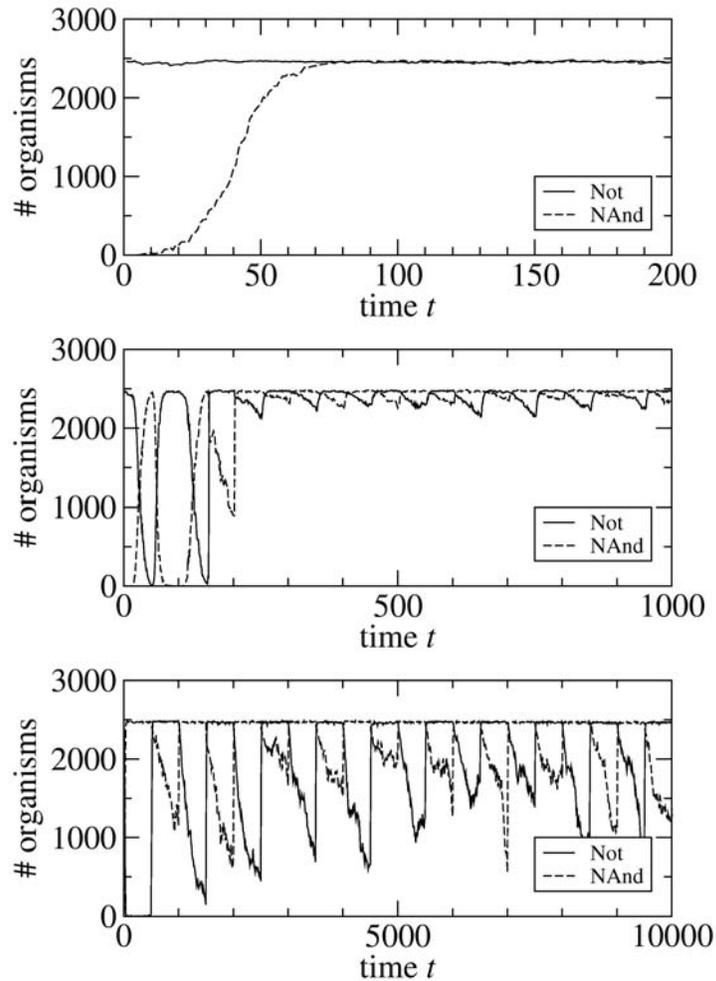


Figure 4. Number of organisms performing tasks Not/NAnd, versus time, for small ($T = 2$, upper panel), medium ($T = 100$, middle panel), and large ($T = 1000$, lower panel) periods, at mutation rate $\mu = 0.1$. Time is measured in generations, where generation 0 is the first generation in the time-dependent landscape (thus, generation 1001 overall).

normal. Rather, it is a double humped distribution that is inaccurately characterized by the mean and variance. We see the same effect in Figure 6d for period lengths $T = 100$ and $T = 1000$.

Second, we notice in Figure 6c and d that there is a time lag between the onset of selection and the maximum of the population's gene frequency. This time lag is known as the *phase shift* [16], and is predicted by quasi-species theory for small and intermediate period lengths. In the extreme, for very short periods, the lag can move the maximum gene frequency into the half period in which the gene is not under selection, as we can see for example in Figure 6c for $T = 10$.

We have estimated the phase shift by determining the phase Φ_{\max} at which the population achieves its maximal gene frequency after $\Phi_0 = 0.5$, that is, after the gene is back under selection. This estimate is shown in Figure 7 for mutation rate $\mu = 0.1$,

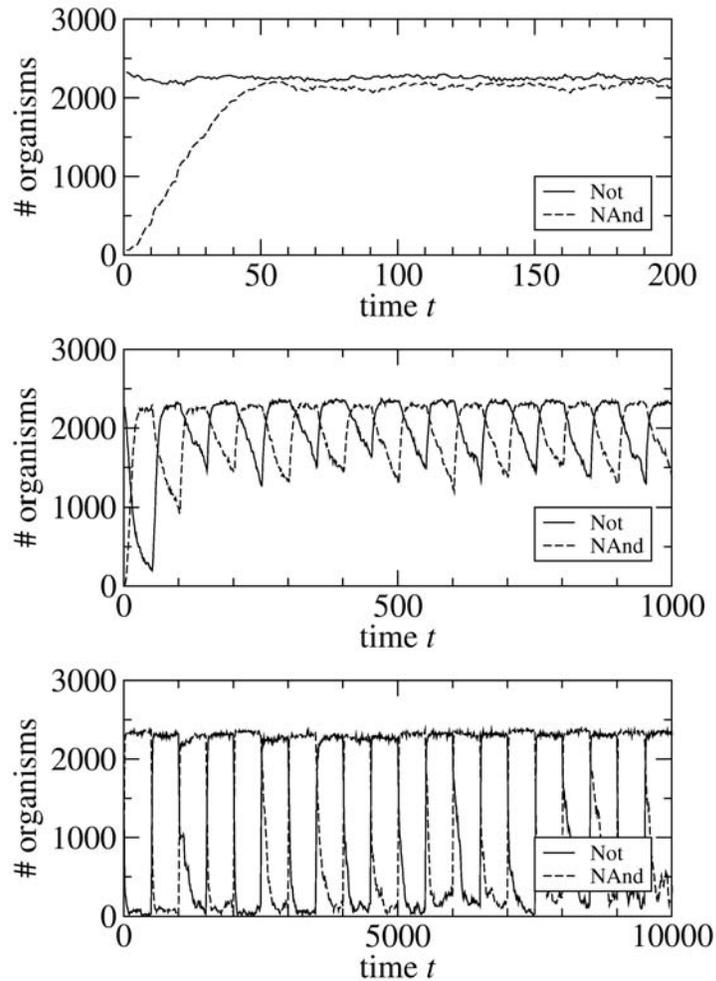


Figure 5. Number of organisms performing tasks **Not**/**NAnd** versus time for small ($T = 2$, upper panel), medium ($T = 100$, middle panel), and large ($T = 1000$, lower panel) periods, at mutation rate $\mu = 1.0$. The notion of time is identical to that of Figure 4.

plotted against the period length. The phase shift decays faster for the **Not**/**NAnd** pair than for the **And**/**OrN** pair, which is in agreement with our earlier assessment that the latter pair is harder to evolve.

Finally, we found that for intermediate period lengths, there exists a selective pressure to rewrite the genome in such a way that the loss of unselected tasks becomes less likely. For example, in Figure 4b, during the first two periods, both **Not** and **NAnd** are lost when they are not selected. However, starting with the third period, now the organisms have rewired their genomes so that loss of either unselected task is minimal. A similar effect can be seen in Figure 8. (The experiment shown in Figure 8 is identical to that of Figure 5b, but we display a longer time frame in Figure 8.) Around generation $t = 1000$, the maximum gene frequency for **NAnd** over one period increases slightly, and the minimum gene frequency over one period increases substantially.

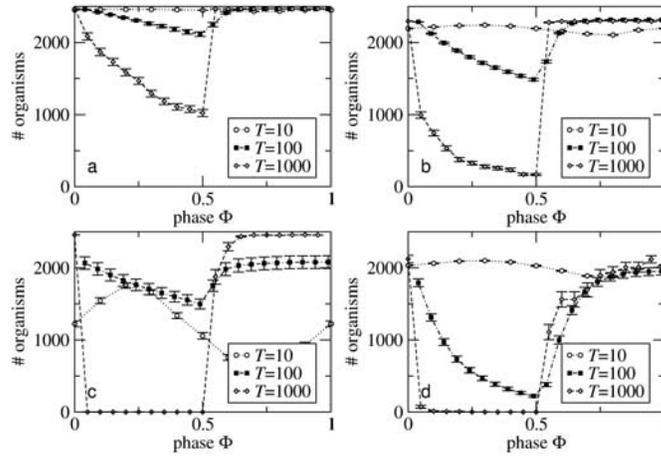


Figure 6. Number of organisms performing Not (a, b) or NAnd (c, d) as a function of the phase $\Phi = t(\text{mod } T)/T$. (a, b) Task frequency for the NAnd/Not pair at $\mu = 0.1$ and $\mu = 1.0$, respectively. (c, d) Task frequency for the And/OrN pair at $\mu = 0.1$ and $\mu = 1.0$, respectively. Frequencies were averaged over all periods in a run. Error bars indicate standard error.

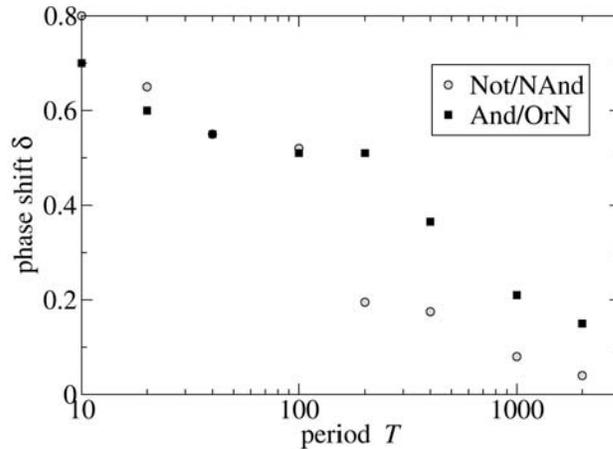


Figure 7. Phase shift δ as a function of period length. The phase shift was estimated as the phase Φ_{max} at which the first task (Not for the Not/NAnd pair, and And for the And/OrN pair) reached its maximum value, minus $\Phi_0 = 0.5$ (at which time the first task starts to be rewarded). For those cases in which $\Phi_{\text{max}} - \Phi_0$ was negative, we used $\delta = 1 + \Phi_{\text{max}} - \Phi_0$.

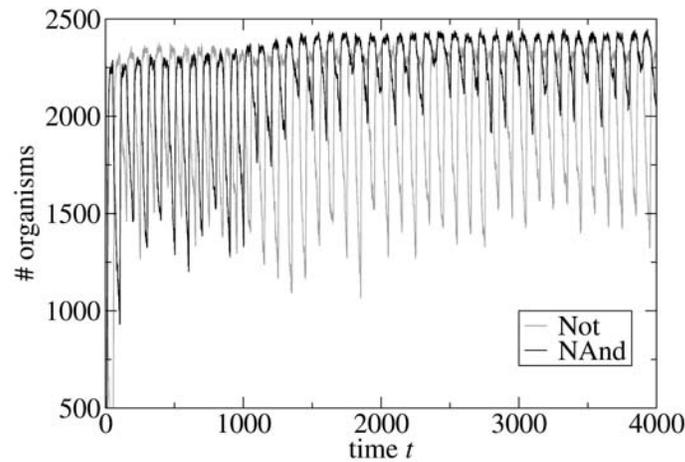


Figure 8. Number of organisms performing tasks Not/NAnd, versus time, for period length $T = 100$, at mutation rate $\mu = 1.0$. Time is measured in generations.

4 Discussion

Overall, our results agree very well with existing theory [14, 29, 16]: For both peak pairs and mutation rates, we found that the population adapts to the average fitness landscape for quickly changing landscapes, and to the particular landscape that is active at a given moment for very slowly changing landscapes.

We found that the phase shift increases up to almost 1 for small T . At first glance, a phase shift this large seems counterintuitive, since a shift larger than 0.5 indicates that genes continue to increase in frequency even if they are not under selection anymore. We can understand this phase shift from the particular way in which organisms acquire fitness in Avida [18]: Fitness in Avida is the ratio of the speed at which an organism's CPU runs (i.e., the organism's *merit*) to the number of instructions the organism executes in one gestation cycle. Organisms inherit their initial merit from their parents, and their merit does not change during one gestation cycle. However, during its gestation cycle, an organism accumulates bonuses that will count towards a new merit value, which is given to both the organism and its offspring at the end of the gestation cycle. It can be shown that this type of maternal fitness determination does not alter the course of evolution in a static fitness landscape [25]. However, it does introduce short delays into the process of adaptation. These short delays become apparent in the dynamics of adaptation in a temporally changing fitness landscape. The moment a task is switched off, organisms that have been carrying out this task still have the accumulated merit that they earned earlier. Once they reproduce, they and their offspring will get this merit value, which means they can continue to reproduce at a high fitness value. As a consequence, the gene frequency can continue to rise for two more generations after the selective pressure has been switched off. We can see this effect very clearly in Figure 6c, for $T = 10$.

We should note that while the effect described in the previous paragraph contributes to the phase shift, it is by no means the sole origin of the phase shift. We can estimate the contribution of this effect to the phase shift by $3/T$, since at most three generations

are affected by the delayed response of organisms to a new bonus structure. If we subtract $3/T$ from the measured phase shift, we find that the phase shift converges to $\delta = 0.5$ for small T (data not shown). This result is intuitively obvious: If the environment changes faster than it takes the population to regrow a task, then the corresponding gene frequency will always have its maximum exactly at the point in time when the task is switched off again.

We found for intermediate period lengths that the organisms followed the changes in the fitness landscape, while at the same time they retained a genetic memory of past environments. By genetic memory, we here mean that the organisms acquired a genetic architecture that minimized loss of the unselected gene: Over time, the organisms were able to rearrange their genomes in such a way that the unselected gene remained in the population in increasingly higher proportions. Presumably, this rearrangement occurred through increased interactions between genes. If the selected and the unselected gene have strong negative epistatic interactions, then mutations in the unselected gene will likely have a negative effect on fitness, and thus be selected against. Similar changes in genetic architecture have been observed as a response to alternating environments in experiments with arboviruses: Vesicular stomatitis virus experienced fitness increases in both cell types when subjected to alternating passages in mammalian and insect cells [17] or different mammalian cells [23]. Adaptation to a single cell type in a constant environment resulted only sometimes in increased fitness in the other cell type. Likewise, alternating passages of eastern equine encephalitis virus in different cell types consistently led to fitness improvements in both cell types, whereas passages in a single cell type did not consistently increase fitness in the other cell type [24, 4].

The possibility that the genetic architecture may change as a result of adaptation to changing environments lies outside the scope of the simple theoretical models published to date [14, 29, 16, 8, 9]. These models are typically built on the basis of a very simple genotype-to-phenotype map without significant interaction between genes. For this reason, we have currently no good theoretical understanding of the conditions under which we can expect to see simultaneous adaptation to different environmental conditions, nor do we understand the genetic mechanisms that enable this adaptation. For example, if the environment remains unchanged for too long, then certainly the population will lose its adaptation to the alternative environment. However, we do not know how long is too long. Second, we mentioned above that loss of the unselected gene is most likely prevented by increased epistatic interactions between genes. However, we do not have evidence for or against this hypothesis, nor is there to date a simple mathematical model that can corroborate or refute it.

Acknowledgments

We would like to thank Allan Drummond for technical assistance, and Chris Adami, Santiago Elena, and Isabel Novella for helpful comments and suggestions. C.O.W. was supported by NSF grant DEB-9981397 to Chris Adami. All data were obtained on a cluster of 160 processors at Michigan State University, supported by the NSF under the aforementioned contract.

References

1. Charlesworth, B. (1976). Recombination modification in a fluctuating environment. *Genetics*, 83, 181–195.
2. Charlesworth, B. (1988). The evolution of mate choice in a fluctuating environment. *Journal of Theoretical Biology*, 130, 191–204.
3. Charlesworth, B. (1993). The evolution of sex and recombination in a varying environment. *Journal of Heredity*, 84, 345–350.

4. Cooper, L. A., & Scott, T. W. (2001). Differential evolution of eastern equine encephalitis virus populations in response to host cell type. *Genetics*, *157*, 1403–1412.
5. Eigen, M. (1971). Self-organization of matter and the evolution of macromolecules. *Naturwissenschaften*, *58*, 465–523.
6. Eigen, M., & Schuster, P. (1979). *The hypercycle—A principle of natural self-organization*. Berlin: Springer-Verlag.
7. Ishii, K., Matsuda, H., Iwasa, Y., & Sasaki, A. (1989). Evolutionary stable mutation rate in a periodically changing environment. *Genetics*, *121*, 163–174.
8. Kamp, C., & Bornholdt, S. (2002). Co-evolution of quasispecies: B-cell mutation rates maximize viral error catastrophes. *Physical Review Letters*, *88*, 068104.
9. Kamp, C., Wilke, C. O., Adami, C., & Bornholdt, S. (2002). Viral evolution under the pressure of an adaptive immune system—Optimal mutation rate for viral escape. *Complexity*, *8*, 28–33.
10. Kimura, M. (1967). On the evolutionary adjustment of spontaneous mutation rates. *Genetical Research*, *9*, 23–24.
11. Laakso, J., Kaitala, V., & Ranta, E. (2001). How does environmental variation translate into biological processes? *Oikos*, *92*, 119–122.
12. Lande, R., & Shannon, S. (1996). The role of genetic variation and population persistence in a changing environment. *Evolution*, *50*, 434–437.
13. Leigh, E. G. (1970). Natural selection and mutability. *American Naturalist*, *104*, 301–305.
14. Nilsson, M., & Snoad, N. (2000). Error thresholds on dynamic fitness landscapes. *Physical Review Letters*, *84*, 191–194.
15. Nilsson, M., & Snoad, N. (2002). Optimal mutation rates in dynamic environments. *Bulletin of Mathematical Biology*, *64*, 1033–1043.
16. Nilsson, M., & Snoad, N. (2002). Quasispecies evolution on a fitness landscape with a fluctuating peak. *Physical Review E*, *65*, 031901.
17. Novella, I. S., Hershey, C. L., Escarmis, C., Domingo, E., & Holland, J. J. (1999). Lack of evolutionary stasis during alternating replication of an arbovirus in insect and mammalian cells. *Journal of Molecular Biology*, *287*, 459–465.
18. Ofria, C., & Wilke, C. O. (2004). Avida: A software platform for research in computational evolutionary biology. *Artificial Life*, *10*, 191–229.
19. Ronnewinkel, C., Wilke, C. O., & Martinetz, T. (2001). Genetic algorithms in time-dependent environments. In Kallel, L., Naudts, B., and Rogers, A. (Eds.), *Theoretical aspects of evolutionary computing, natural computing* (pp. 261–285). Springer-Verlag.
20. Sniegowski, P. D., Gerrish, P. J., & Lenski, R. E. (1997). Evolution of high mutation rates in experimental populations of *E. coli*. *Nature*, *387*, 703–705.
21. Taddei, F., Radman, M., Maynard-Smith, J., Toupance, B., Gouyon, P. H., & Godelle, B. (1997). Role of mutator alleles in adaptive evolution. *Nature*, *387*, 700–702.
22. Travis, J. M. J., & Travis, E. R. (2002). Mutator dynamics in fluctuating environments. *Proceedings of the Royal Society of London Series B*, *269*, 591–597.
23. Turner, P. E., & Elena, S. F. (2000). Cost of host radiation in an RNA virus. *Genetics*, *156*, 1465–1470.
24. Weaver, S. C., Brault, A. C., Kang, W., & Holland, J. J. (1999). Genetic and fitness changes accompanying adaptation of an arbovirus to vertebrate and invertebrate cells. *Journal of Virology*, *73*, 4316–4326.
25. Wilke, C. O. (2002). Maternal effects in molecular evolution. *Physical Review Letters*, *88*, 078101.
26. Wilke, C. O., & Adami, C. (2002). The biology of digital organisms. *Trends in Ecology & Evolution*, *17*, 528–532.

27. Wilke, C. O., & Ronnewinkel, C. (2001). Dynamic fitness landscapes: Expansions for small mutation rates. *Physica A*, *290*, 475–490.
28. Wilke, C. O., Ronnewinkel, C., & Martinetz, T. (1999). Molecular evolution in time-dependent environments. *Lecture Notes in Artificial Intelligence*, *1674*, 417–421.
29. Wilke, C. O., Ronnewinkel, C., & Martinetz, T. (2001). Dynamic fitness landscapes in molecular evolution. *Physics Reports*, *349*, 395–446.
30. Wilke, C. O., Wang, J. L., Ofria, C., Lenski, R. E., & Adami, C. (2001). Evolution of digital organisms at high mutation rates leads to survival of the flattest. *Nature*, *412*, 331–333.