

# Influence of chance, history and adaptation on evolution in *Digitalia*

Daniel Wagenaar and Chris Adami

California Institute of Technology, Pasadena, CA 91125

## Abstract

We evolved multiple clones of populations of *Digitalia*, a type of digital organism, to study the effects of chance, history, and adaptation in evolution. We show that clones adapted to a specific environment can adapt to new environments quickly and efficiently, although their history remains a significant factor in their fitness. Adaptation is most significant (and the effects of history less so) if the old and new environments are dissimilar. For more similar environments, adaptation is slower while history is more prominent. For both similar and dissimilar transfer environments, populations quickly lose the capability to perform computations (the analog of beneficial chemical reactions) that are no longer rewarded in the new environment. Populations that developed few computational “genes” in their original environment were unable to acquire them in the new environment.

## Introduction

One of the central tenets of standard evolutionary theory is that characteristics of evolved populations can be explained by the process of adaptation, and thus that phenotypic differences, for the most part, have an adaptive value. That all of biological diversity is due to adaptation has been challenged in modern expositions of Darwinian theory. Kimura (1983), for example, proposes that there is a strong component of chance in evolution, while Gould and Lewontin (1979) stress the importance of history and contingency. Effects of chance are usually due to genetic drift and random mutations without value to the organism. History can become important if certain genetic changes (of adaptive value in the past) constrain or promote some evolutionary outcomes over others. To disentangle these effects, Gould (1989) has proposed to “replay the tape” of evolution to test its repeatability. Travisano, Mongold, Bennett and Lenski (1995) were the first to perform a rigorous experiment of this sort, albeit on a shorter timescale and with *E. coli* bacteria adapting to simple, artificial environments. The trait undergoing evolution in these experiments was fitness (measured as the Malthusian parameter). As a control, bacterial size (which in these environments is selectively neutral) was also monitored. This study found

that adaptation contributed most significantly to the evolutionary changes, often resulting in convergent evolution of fitness. The evolution of bacterial size, on the contrary, was influenced much more strongly by chance and history, as expected.

To date, this experiment has not been repeated with any other organism, nor could it be determined whether the relative effects of chance, history and adaptation are constant throughout evolutionary time. The advent of digital organisms opens the door to experiments much more alike to “replaying the tape of life”, and also to test the validity of Travisano et al.’s results across organisms.

Digital organisms have been studied in a variety of experiments pertaining to evolution. They offer a tantalizing glimpse into the characteristics of living systems that do not share any ancestry with biochemical life on earth. These organisms are self-replicating strands of computer code competing for resources in a user-defined environment (the artificial “Petri” dish) within a computer’s memory. Our *Digitalia* live in a world created and controlled through the AVIDA software developed at Caltech (Ofria 1999, Adami 1998, Lenski et al. 1999).

In the experiments reported here, we essentially follow the protocol of Travisano et al., measure similar phenotypic characteristics, and perform the same statistical analysis on the data obtained. However, due to the ease of these experiments with *Digitalia* (as compared to *E. coli* experiments), we are able to collect data much more frequently, allowing the observation of *changes* in the relative importance of chance, history, and adaptation as a function of time. While Travisano et al. used bacteria which had adapted to using glucose as their primary sugar, and studied their re-adaptation to a maltose environment, populations of *Digitalia* are transferred to new environments which award differing computational tasks. As the replication speed of *Digitalia* is mainly due to their computational prowess on random numbers in their environment (see Adami 1998, Ofria 1999) we can change landscapes simply by changing the set of computations which result in extra CPU time for the organism that achieves them. Thus, we adapt *Digitalia* to one landscape first, then transfer them to another and mon-

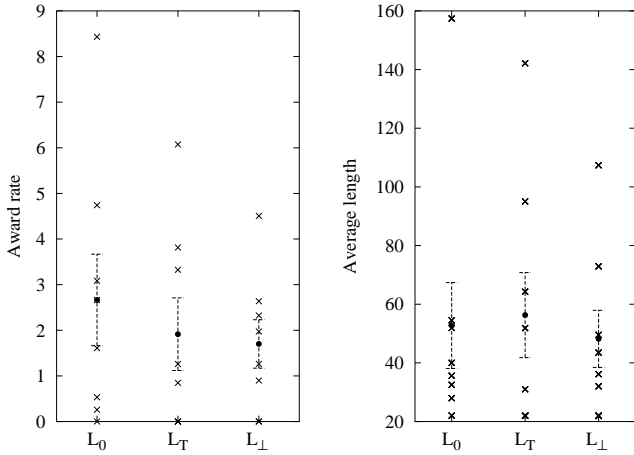


Figure 1: Award rates and average genome lengths attained by populations in their native landscape after 20,000 updates. Each cross represents a population. For the purpose of comparison across landscapes, in this plot (only), all  $c_t^{(L)}$ 's have been set to unity. Averages and their uncertainties are indicated by filled circles with errorbars.

itor their re-adaptation. “Immigrant” populations are expected to have a lower fitness in their new environment than populations that are native to that environment. We investigate how well the immigrants are able to recover this initial disadvantage, and whether in their new environment they carry with them long term effects from their evolutionary history preceding the transfer. In addition to fitness we also observe genome length, a phenotypic trait which we expect to be selectively neutral as far as the differences between our landscapes is concerned. This variable thus substitutes for the role played by bacterial size in Travisano et al.

## Methods

We studied the effects of adaptation and history by evolving populations in one of three distinct environments.  $L_0$  is the standard landscape used in most experiments to-date (Ofria 1999) rewarding a total of 76 different two- and three-input logical operations on random numbers. We split these logical operations into two orthogonal sets which are used to define the landscapes  $L_T$  and  $L_\perp$ <sup>1</sup>. We test the significance of history and adaptation in transfers  $L_0 \rightarrow L_T$  (similar landscapes) as well as  $L_\perp \rightarrow L_T$  (transfer to a dissimilar one). In each experiment, we evolve eight separate populations in their environment of origin until they are well adapted, after which the entire populations are cloned five-fold and propagated in their transfer landscape. Genome length

<sup>1</sup>A definition of the landscapes in terms of their rewarded tasks can be found in the Appendix.

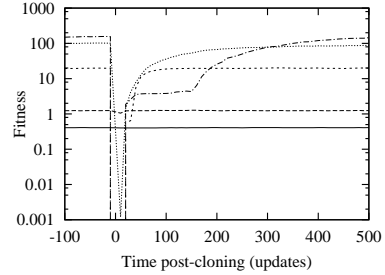


Figure 2: Impact on measured fitness of five cloned populations after re-insertion into their *original* landscape. In this re-transfer control, no adaptation occurs while the original fitness is recovered by AVIDA.

and parameters reflecting the average fitness of the populations are measured at various stages during evolution in the new environment, in order to be compared to the values at the time of the transfer.

To ascertain that the observed effects are not due to peculiarities of the landscapes we constructed, we check that  $L_T$ ,  $L_\perp$  and  $L_0$  are equally challenging. The comparison of “award rates” (a measure reflecting fitness introduced below) for populations adapting to  $L_T$ ,  $L_\perp$  and  $L_0$  depicted in Fig. 1 shows that this is so.

For *Digitalia* as for bacteria, average fitness can usually be measured directly and used as an indicator of the extent of adaptation. Here, we had to forgo this direct approach because the fitness of the transferred population is not well reflected in the measurements right after the transfer, simply because the AVIDA software cannot, at present, accurately monitor a genome’s performance in a new environment until at least one replication cycle has been completed. As a consequence, the average fitness is incorrectly measured for a few hundred updates<sup>2</sup> (see Figure 2). This measurement error masks adaptive events occurring early after transfer and as a consequence would severely compromise the analysis.

Instead, we study the *award rate*, a variable closely related to average fitness. The award rate  $\mathcal{A}^{(L)}(p)$  of a population  $p$  in a landscape  $L$  is defined as

$$\mathcal{A}^{(L)}(p) = \sum_t c_t^{(L)} \mathcal{F}_t(p),$$

where  $\mathcal{F}_t(p)$  is the fraction of creatures in  $p$  that perform task  $t$ , and the coefficients  $c_t^{(L)}$  are unity for tasks that are rewarded in landscape  $L$ , and zero for all other tasks. As opposed to average measured fitness, the award rate is not significantly affected by the cloning operation.

Following Travisano et al. (1995), we plot the *derived* (or adapted) value of the characteristic versus the *an-*

<sup>2</sup>Time is measured in arbitrary units called updates; every update represents the execution of an average of 30 instructions per individual in the population. A typical generation takes 5-10 updates.

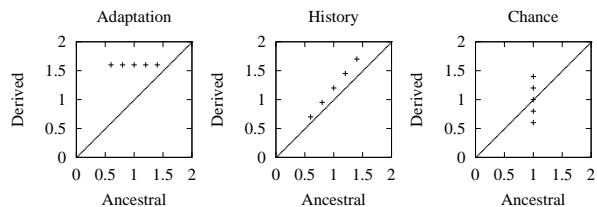


Figure 3: Derived versus ancestral values of hypothetical traits, the evolution of which is chiefly determined by adaptation (left panel), history (middle panel) and chance (right panel). Note that the effects of chance are demonstrated for a set of clones of a single ancestral genotype, whereas adaptation and history are illustrated for several independent ancestors. Adapted from Travisano et al. (1995).

central one at the time of transfer, in order to study the effects of chance, history and adaptation on the evolution of a population’s characteristics (see Fig. 3).

Eight populations were evolved for 20,000 updates in each of the landscapes  $L_0$ ,  $L_T$  and  $L_\perp$ , and cloned<sup>3</sup>. In both transfer experiments, cloned populations were propagated for an additional 10,000 updates of re-adaptation.

The relative contributions of adaptation, history, and chance to the evolution of traits such as fitness and genome length can be disentangled by studying the variance of the respective observable. A nested anova (see Sokal and Rohlf 1995) is used to determine what fraction of the variance (after evolution) should be attributed to the elements of history and chance. The contribution of adaptation is obtained from the average difference between derived values and ancestral values.

Award rate and genome length for each population are sampled over a range of 30 updates every 1,000 updates. The spread within these samples is used as the measurement error. This does not reflect the intra-population spread of the parameter under consideration, but only their natural short-time variability. Populations that had award rates less than 0.002 (as measured in their old landscapes) at the time of the transfer, were excluded from the analysis<sup>4</sup>. For the transfer  $L_\perp \rightarrow L_T$ , two out of eight sets of clones were excluded, while only one set of clones was excluded for the transfer  $L_0 \rightarrow L_T$ . None of

<sup>3</sup>We checked that propagation for an additional 10,000 updates in their native landscapes produced no significant further increase of award rate, so the populations can be said to be at or near equilibrium in their particular environments.

<sup>4</sup>Due to a special feature in the present physics of the AVIDA world, genomes can evolve that cannot alter their size, preventing any further adaptation. These cases can be considered anomalous as adaptation cannot be measured. We plan to alter the physics of replication in future versions of AVIDA to avoid such contaminations.

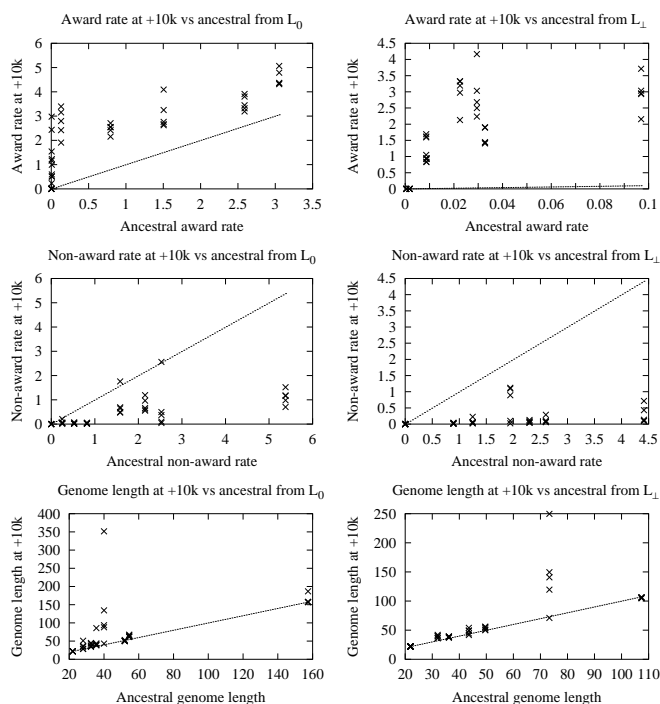


Figure 4: Award rates (top) and average genome length (bottom) for populations transferred from landscape  $L_0$  and  $L_\perp$  to landscape  $L_T$ : values at 10,000 updates after the transfer plotted against values at the time of the transfer. Middle panels show the ‘non-award rate’.

the excluded populations attained a derived award rate significantly above 0.002 after transfer and propagation. The lowest recorded derived award rate for any of the other populations was  $0.23 \pm 0.05$ .

## Results

We found that adaptation is the dominant component of evolutionary changes of fitness in *Digitaria*, mirroring the results of Travisano et al. obtained for *E. coli*. This can be seen even without statistical analysis by noting that the award rates in Fig. 4 are consistently higher after adaptation to a new landscape than upon transfer, independently from whether the transfer landscapes are similar to the ancestral one or not. Genome length did not change significantly as a result of re-adaptation, confirming that this trait is selectively neutral. Plotting the relative contributions of adaptation, chance, and history as a function of time (Figures 5, 6) reveals that adaptation of fitness is always more important than chance. It is dominated by history at first, but ultimately becomes the principal component of evolutionary change. Conversely, chance and history were dominant in the evolution of genome length. This indicates that genome length indeed does not discriminate between the different landscapes, i.e., that the amount of information in the environments is similar.

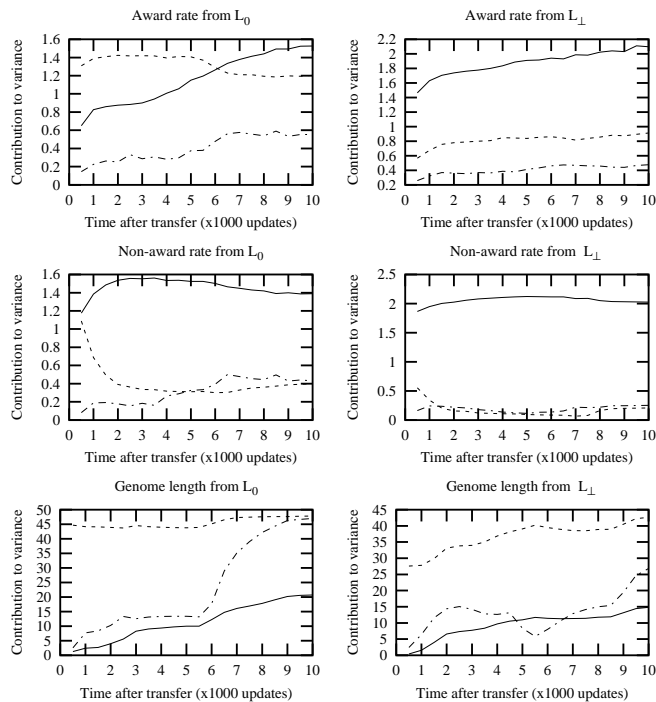


Figure 5: Contributions of adaptation (lines), history (dashes) and chance (dash-dotted) to variance in award rate (top), non-award rate (middle) and genome length (bottom) in populations transferred from landscape  $L_0$  (left) and  $L_{\perp}$  (right) to landscape  $L_T$ .

The adaptation effects on award rate and ‘non-award’ rate (the rate at which tasks are performed which are *not* rewarded in the landscape) mostly take place in the first 1,000 updates (see Figures 5, 6). After 10,000 updates the award rates of the transferred clones do not differ significantly from the award rates of populations that evolved *de novo* in landscape  $L_T$  for 20,000+10,000 updates, indicating convergent evolution.

Note that the dominance of adaptation and history over chance after 10,000 updates is statistically highly significant (see confidence limits in Figure 7), whereas the dominance of adaptation over history is not. The significance of this dominance should be ascertained in experiments in which the period of adaptation is extended.

## Conclusions and Outlook

We have found that populations of *Digitalia* propagated within an artificial world created by the AVIDA software are able to adapt to a new environment even though they are already well adapted to another environment, and even if the new landscape is orthogonal to the old one in terms of rewarded behavior. The fitness of organisms in the new landscape continues to be strongly influenced by their history in the old landscape, more so

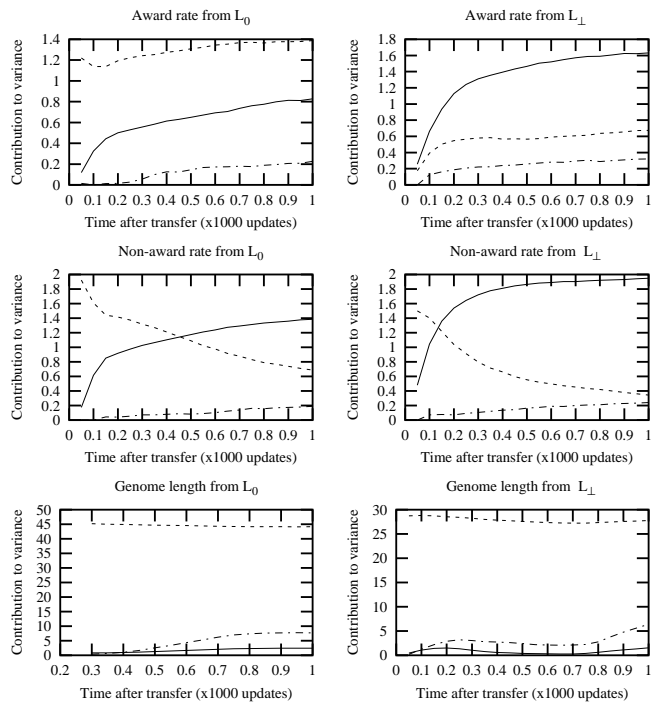


Figure 6: Expanded view of the early history (the first 1,000 updates) of adaptation shown in Figure 5. Award rates and genome lengths for updates earlier than 200 updates may be affected by the transfer (although no long term effects persist).

than in the bacterial experiments with *E. coli* referred to above, and this influence does not appear to decrease after 10,000 updates. However, the effect of adaptation was still increasing when each of the experiments was terminated. This suggests that follow-up experiments with longer adaptation in the transfer environment might be required to completely assess the influence of adaptation. As a significant part of the adaptation was found to take place in the first 1,000 updates after the transfer, it is imperative to repeat these experiments studying the actual average fitness of the population instead of the award rate, and verify the present results. Experiments of this nature with an updated version of AVIDA are planned for the near future (Smith and Wagenaar 2000).

The high speed of adaptation post transfer indicates that populations are able to change the tasks they perform, and so as it were rearrange their metabolism which provides the energy for their replication, by simple mutations only. This suggests that, although the organisms were highly adapted to their native landscape, they retained a significant amount of plasticity which allowed them to thrive in changed environments without much effort. The plasticity of genes in molecular biology is the subject of much discussion (see, e.g., Via et al, 1995) in particular with the advent of whole genomes in the age

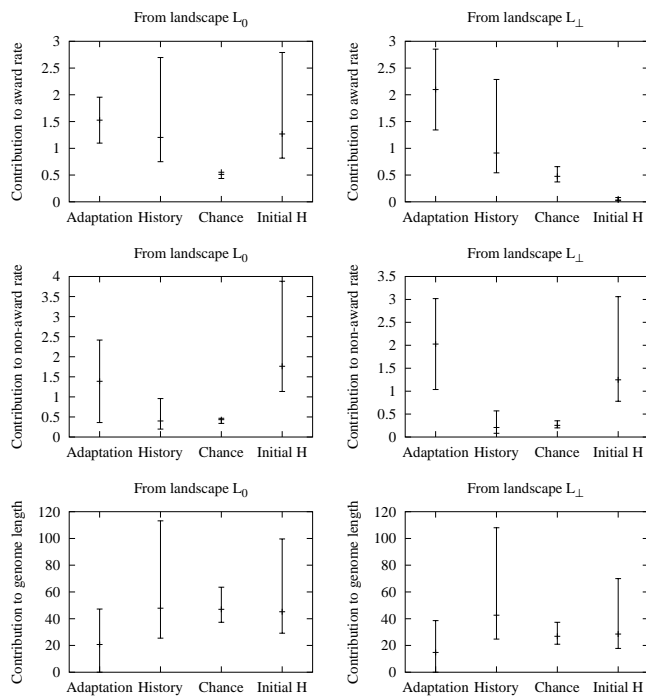


Figure 7: Confidence limits for the relative contributions of history, chance and adaptation at 10,000 updates post transfer. ‘Initial H’ is the initial spread of the variable due to history in the original landscape. This value is extremely low for the  $L_{\perp} \rightarrow L_T$  award rate due to the very low initial fitness of these populations, as landscapes  $L_T$  and  $L_{\perp}$  are orthogonal.

of bioinformatics. That it can also be observed in the evolution of digital organisms suggests that it may be a universal feature of evolution.

D. W. wishes to thank Justin Smith for useful discussions. This material is based upon work supported by the NSF under Award No. DEB-9981397.

## Appendix – Avida parameters

Experiments reported here were performed with version 1.3.1 of the AVIDA software, which can be obtained from <http://www.krl.caltech.edu/avida>. We used populations of  $60 \times 60$  organisms, with standard instruction set, and removal of the oldest as birth method. Genomes were subject to a copy mutation rate of 0.7% per instruction copied, and an insert/delete probability of 5% per generation. Rewards for performing logical functions were set as follows. Multiplicative merits for two input logic functions:

Operation	$L_0$		$L_T$		$L_{\perp}$	
	1st	2nd	1st	2nd	1st	2nd
$A \wedge B$	1.2	1.1	-	-	1.5	1.3
$A \vee B$	1.25	1.1	-	-	1.5	1.3
$A \vee \overline{B}$	1.2	1.1	1.5	1.3	-	-
$A \wedge \overline{B}$	1.25	1.1	1.5	1.3	-	-
$\overline{A \vee B}$	1.3	1.1	-	-	1.5	1.3
$\overline{A \wedge B}$	1.15	1.1	-	-	1.3	1.2
$A \neq B$	1.5	1.1	-	-	1.8	1.5
$A = B$	1.5	1.1	1.8	1.8	-	-

Multiplicative merits for three input logic functions:

Nos	$L_0$		$L_T$		$L_{\perp}$	
	1st	2nd	1st	2nd	1st	2nd
odd	1.5	1.1	-	-	2.2	1.5
even	1.5	1.1	2.0	1.4	-	-

Note that merits for landscape  $L_{\perp}$  are consistently higher than for landscape  $L_T$ . This was done because a variant of landscape  $L_{\perp}$  with lower merits turned out to be unlearnable. The merits for the trivial (one-input) operations were set to 1.05 for all landscapes. For calculating the award rates, however, only the two and three input logic functions are taken into account. Any non-unity merit caused that task to be counted with a weight  $c_t = 1$ .

## References

- Adami, C., 1998. *Introduction to Artificial Life*. Santa Clara: TELOS Springer-Verlag.
- Gould, S.J., 1989. *Wonderful Life: The Burgess Shale and the Nature of History*. New York: Norton.
- Gould, S.J. and Lewontin, R.C., 1979. The spandrels of San Marco and the panglossian paradigm: A critique of the adaptationist programme. *Proc. R. Soc. London Ser. B* **205**:581–598.
- Kimura, M., 1983. *The Neutral Theory of Molecular Evolution*. Cambridge: Cambridge University Press.
- Lenski, R.E., Ofria, C., Collier, T. C., and Adami, C., 1999. Genome complexity, robustness, and genetic interaction in digital organisms. *Nature* **400**: 661–663.
- Ofria, C. A., 1999. *Evolution of Genetic Codes*, Ph.D. thesis, Caltech, Pasadena, CA 91125.
- Smith, J., and Wagenaar, D. A., 2000. Early stages of evolution of newly cloned digital organisms. Work in progress.
- Sokal, R. R. and Rohlf, F. J., 1995. *Biometry, 3rd ed.* New York: W. H. Freeman.
- Travisano, M., Mongold, J. A., Bennett, A. F., and Lenski, R. E., 1995. Experimental test of the roles of adaptation, chance and history in evolution. *Science* **267**: 87–90.
- Via, S. et al., 1995. Adaptive phenotypic plasticity—Consensus and controversy. *Trends Ecol. Evol.* **10**: 212–217.