Finger-painting Fitness Landscapes: An Interactive Tool for Exploring Complex Evolutionary Dynamics.

Luis Zaman[1,2,3], Charles Ofria[1,2], Richard E. Lenski[2,3]

[1]Department of Computer Science and Engineering
[2]BEACON Center for the Study of Evolution in Action
[3]Ecology, Evolutionary Biology & Behavior Program
Michigan State University, East Lansing, MI USA
{zamanlui, ofria, lenski}@msu.edu

Abstract

Evolution involves only a few simple processes, yet the resulting dynamics are surprisingly rich and complex. Sewall Wright developed the metaphor of fitness landscapes to provide deeper insight into the complex workings of evolution. Here we extend that metaphor by visualizing in real time the dynamic processes that drive evolution. We allow viewers to construct fitness landscapes interactively while also varying key parameters including population size, mutation effect size, mode of reproduction (asexual or sexual), and density-dependent selection. This application is both mechanistic and visual, and it thereby allows the active exploration of evolutionary processes. We walk the reader through several exercises including both simple activities potentially suitable for education and examples of deeply conceptual topics that remain the focus of current research in evolutionary biology.

Introduction

Sewall Wright first depicted fitness landscapes in 1932 as a contour plot relating two genetic axes with hills and valleys of fitness (Wright, 1932), and these landscapes are still pervasive in evolutionary biology today. Even at this inception, Wright understood the oversimplifications necessary to depict genetic space in so few dimensions. However, the insights and intuition this visual metaphor has brought to evolutionary thinking are great (Wright, 1988; Gavrilets, 2004). In fact, one of the most important questions in evolution concerns how populations move from local optima to higher (potentially global) optima – a question in which fitness landscapes are central (Pigliucci and Kaplan, 2006). Substantial work has been done to address this question, from appealing to the unintuitive geometry of high-dimensional spaces by Fisher (Whitlock et al., 1995; Orr, 1998) to considering landscapes with mostly neutral mutations by Kimura (1983).

Fitness landscapes play a substantial role in evolutionary computation as well as biology, though the relevance of computational landscapes is less debated because evaluation functions are sufficient to describe fitness surfaces for most optimization problems. For example, Langdon investigated the structure of fitness landscapes for some canonical evolutionary computation functions such as XOR (Langdon and Poli, 1999; Langdon, 1999). Just as evolutionary biology informs computation, sometimes computation can shed light on biology. Kashtan et al. (2007) showed that changing environments in a digital system gave populations access to peaks they otherwise could not explore. Experiments in Avida, an artificial life platform, demonstrated a "survival of the flattest" effect: at high mutation rates, populations evolved to lower and flatter rather than higher and steeper regions in the fitness landscape (Wilke et al., 2001; Wilke and Adami, 2003).

While the landscape metaphor holds a prominent place in evolutionary thinking, it is not without critics. One criticism concerns the metaphor's multiple forms: one describing individual fitness as a function of genotypes, another as a function of phenotypes (Simpson, 1953), and yet another describing a population's mean fitness as a function of its genetic structure (Pigliucci and Kaplan, 2006). Except in a few special cases, the axes are not rigorously defined, but rather depict some sort of distance between types.

A related set of criticisms of the fitness landscape metaphor concern the lack of rigorous mathematical formalism (Provine, 1989). In Wright's defense, the metaphor was meant to hide the mathematics necessary for describing evolution in massively multi-dimensional spaces, while providing an intuitive framework for considering the various possible outcomes. Gavrilets (2004) distinguishes the mathematical fitness landscape as a high-dimensional formal construct, but he still must show them in two or three dimensions. Rigorous mathematics are necessary for advancing theory in high-dimensional landscapes, but the formalisms may provide little intuition about the evolutionary process. Perhaps this is why, despite the criticisms, depictions of fitness landscapes usually reflect Wright's original form. Many fundamental concepts in evolution can be illuminated using so simple a metaphor.

Depicting whole populations evolving on fitness landscapes is even more complicated; they are often shown as an abstract cloud moving up a peak. Numerical simulations and other analytical methods are generally required for
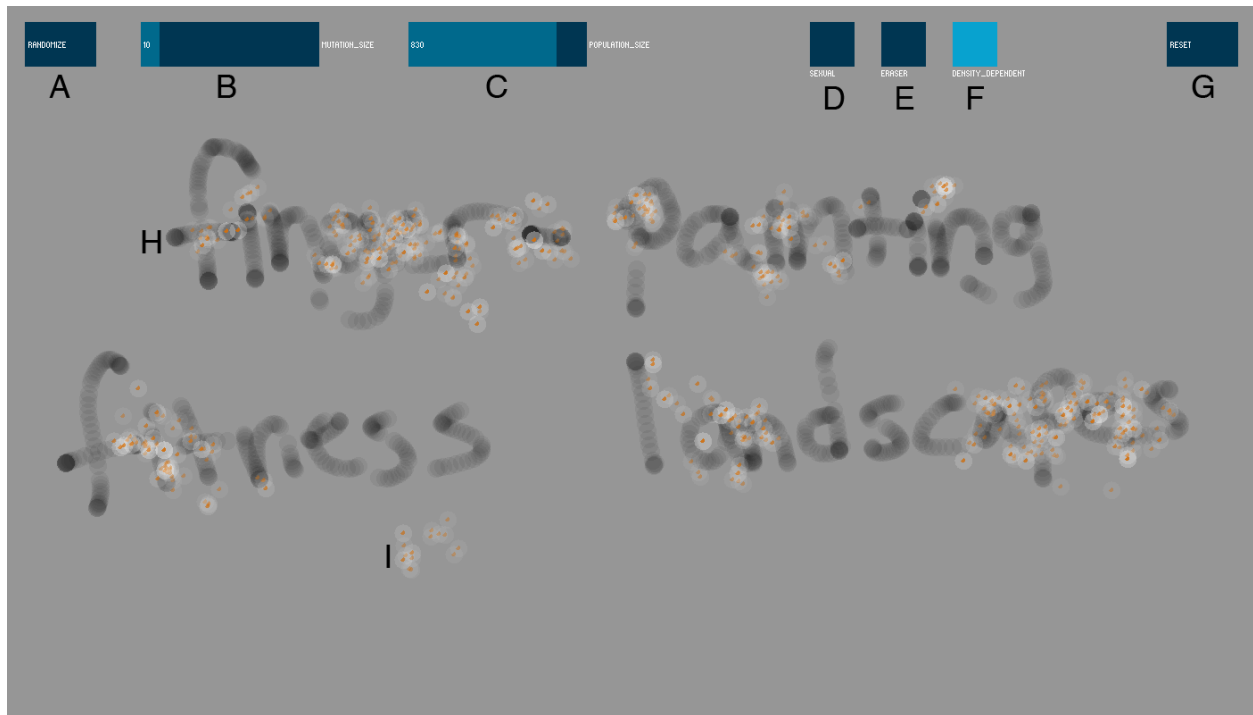
Figure 1: The full screen application view. (A) A button that randomizes each individual's genotype, spreading them out across the fitness surface. (B) A slider that controls the size of mutational effects, which are drawn from a uniform distribution and are applied randomly to one of the two genetic dimensions every time an offspring is produced. (C) A slider that controls population size. (D) An asexual-sexual toggle that, when enabled, causes two parent organisms to produce offspring with an averaged (in both dimensions) genotype for 30% of reproductive events (70% of reproductive events remain asexual). (E) An eraser toggle that switches from the drawing mode to an erasing mode. (F) A toggle for density-dependence, shown enabled, allows the viewer to explore how deformations of the landscape caused by the organisms (e.g., by depleting resources) affect evolutionary dynamics. (G) A reset button that removes all painted regions and randomizes the genotypes. (H) A painted region that represents the fitness surface, where darker areas depict higher fitness. Repeated or slower strokes increase the darkness of touched regions, producing higher fitness peaks. (I) Individuals are depicted as small orange squares, shown here with density-dependence enabled, which suppresses the fitness surface immediately surrounding them, thus lightening the region.

deeper insight about the typical and exceptional paths that populations may take. However, these methods require substantial time and expertise to master, whereas the intuitive finger-painting system that we present provides a simple, fast heuristic tool for interactive discovery. Thus, despite the limitations of Wright's two-dimensional genetic space, it is the most accessible form of fitness landscapes for visualizing the evolutionary process in action and, as such, the representation that we chose to extend. We note, however, that mutations in our system can move genotypes various distances (not single uniform steps) from their progenitors, similar to mutations on phenotypic landscapes.

Interactive System and Touchscreen Display

We built an interactive system that combines visualization and simulation, allowing the user to construct and modify a fitness landscape on which a population evolves in real-time.

Developed for a touchscreen interface, the user can "finger-paint" diverse fitness landscapes. Each stroke slightly darkens the surface, and regions become even darker with additional strokes to the same area. In the visualization, darker regions depict higher fitness levels. The light-gray background regions represent a baseline fitness level, while the maximum fitness level is $\sim 70\%$ higher than the baseline; it requires ~ 100 strokes of a given spot to produce the maximum fitness. These and many other details of the implementation can be changed by modifying the underlying program, but they are not subject to change by the user. However, in addition to finger-painting the fitness landscape, the user can vary several key aspects of the simulation by using toggles and sliders. There is an eraser toggle that, when activated, causes additional touches to restore the corresponding regions to the low baseline fitness. There is also a reset button that allows the user to erase all painted areas. These sim-

ple options let users quickly build complex landscapes that include such important features as hills, valleys, ridges, and plateaus.

In addition to the landscape, the system also displays an evolving population of organisms, with each individual genotype located on the fitness landscape and shown as a semi-transparent orange square. Each genotype has two integer values that provide its coordinates on the drawing surface. The program simulates evolution using continuous rounds of tournament selection. In each round, five individuals are randomly sampled from the population and one individual reproduces with a probability that is determined by its fitness as a proportion of the sum of the five fitness values; each fitness value corresponds to the darkness of the fitness surface at the point where the individual sits. When an individual reproduces, it replaces a randomly chosen organism, thus maintaining a constant population size. The population size can be varied by the user, both before and during a given session, using a slider on the interface.

Offspring are mutated along one randomly chosen dimension, with an offset to the parent's coordinate drawn from a random uniform distribution centered at zero. The range of the distribution is determined by the mutation effect size, which can be adjusted interactively by using another slider. Two additional toggles allow the user to vary reproductive mode (asexual versus sexual) and ecological interactions (negative density-dependent effects). Sexual recombination, when enabled, occurs with a 30% probability at every reproduction event. When recombination is triggered, two parents are chosen by tournament selection, and an offspring is produced by averaging the parents' genotypes in both dimensions. When density-dependent selection is implemented, the fitness of each individual is reduced as it interacts with an increasing number of other individuals. In our simulation, this density-dependence acts over local regions of the fitness landscape rather than globally across the whole population. This local interaction may occur if, for example, different fitness peaks represent different resources that can be drawn down by some genotypes but not others. Thus, the more individuals located in a particular region of the fitness landscape, the lower each individual's fitness will be. We show this dynamic on the screen by lightening the surface (lowering the fitness) in a small circular region around each individual; the surface becomes progressively lighter in regions with higher densities of organisms. We calculate the density of a region using a hidden layer that specifies the radius of density-dependent effects and allows the program to compute quickly the relevant fitness modifiers. Individual fitness is calculated as $(1 - Density) * PaintedFitness$, where $Density$ is a scaled value between zero and one that represents how depressed the landscape is at a given position. When density-dependence is disabled, $Density$ is always set to zero.

There are, of course, important limitations to our sys-

tem, including the representation of all genotypes in a two-dimensional space. In that respect, our system suffers from the same defect as Wright's metaphorical fitness landscapes, as we discussed in the introduction. On the other hand, we have brought this important metaphor to life by allowing the user to paint endless forms of landscapes and then watch the process of evolution in action on the fitness surfaces. Moreover, the user can alter features of the landscape and manipulate key variables even as evolution proceeds.

In the next section, we describe and illustrate several exercises that can be performed using our program. The source code can be downloaded from http://bit.ly/xn8isR. Execution requires the Processing Development Environment, which is available from http://processing.org/. Additionally, a limited version of the system is viewable in some browsers at http://bit.ly/zJ7B4N.

Exercises for the Reader

The finger-painting application is intended to help the user gain intuition about the dynamics of evolution on fitness landscapes. To that end, we outline below four "exercises for the reader" that span a wide range of evolutionary principles. We begin with depictions of two basic and well-known concepts perhaps appropriate for educational activities: the hill-climbing process driven by natural selection; and the potential for random drift to allow small populations to cross fitness valleys and thereby discover other nearby fitness peaks. We then present two more exercises that illustrate areas of active research: the role of density-dependent effects in flattening the fitness landscape and thus promoting diversity; and how high mutation pressure can favor organisms that occupy flatter, rather than higher, regions of the fitness landscape.

Hill Climbing

Natural selection reflects disproportionate reproduction by *individuals* with high fitness. In the context of fitness landscapes, natural selection is often described as a hill-climbing process, whereby the *population* moves from regions of lower to higher fitness. Despite its intuitive simplicity for those familiar with the basic ideas, there are confusing aspects of the hill-climbing metaphor, especially the important distinction between the unguided behavior of individuals and the systematic advance of the entire population up a local fitness peak. By seeing the process of individuals producing more or fewer offspring based on their fitness levels, and the resulting hill-climbing effect in the population, the user may develop a mechanistic understanding of evolution by natural selection.

To illustrate this process, start by gently touching the screen to create a low (light gray) peak on the fitness surface, as shown in Figure 2 A. After the population has converged on this peak (pressing the Randomize button on the screen will re-disperse the population if necessary), begin drawing
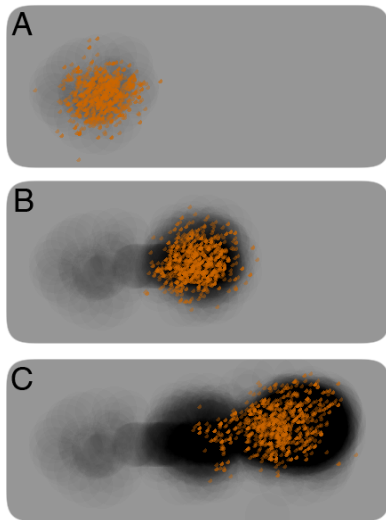
Figure 2: Demonstration of hill-climbing dynamic. (A) Begin with the population converged on a fairly low fitness peak. (B and C) Add progressively darker regions of higher fitness, taking care so that these new peaks are not separated by a wide fitness valley, and watch the population move uphill as more and more individuals occupy the higher regions of the landscape.

a progressively darker region adjacent to the initial peak. If the regions overlap, you will see some new individuals - the product of reproduction and mutation - near the new peak. And because the individuals that are nearer to the new peak have higher fitness, they will reproduce more offspring, so that the population as a whole climbs toward the new peak. If the initial low peak and new higher peak do not overlap, then you can draw a bridge of intermediate fitness that connects them, as shown in Figure 2 B. You can continue this process by drawing additional nearby peaks that are progressively darker and thereby observe the hill-climbing dynamics of an evolving population, as shown in Figure 2 C.

You might then re-start the simulation by pressing the reset button. After proceeding as before through one or two rounds of building adjacent peaks, you can then draw an even higher peak but at a large distance from the other peaks (not shown). You should see that the population does not immediately (if ever) climb that distant peak, despite its high fitness. The different population behavior with respect to connected and disconnected regions of the fitness landscapes shows that evolution finds local fitness peaks more readily than global ones.

Small Populations and Drift

Evolution involves the interplay of several underlying processes. Natural selection reflects the differences in *expected* reproductive success of individuals based on their genotypes

and their fit to the environment. In the context of our application, an individual's expected reproductive success is proportional to the darkness of its location in the fitness landscape. But each individual's *realized* reproductive success also depends on chance. In our application, the tournament selection probabilistically favors more fit individuals but does not guarantee that the most fit will reproduce and, moreover, any individual may be eliminated at random whenever another individual reproduces. In evolutionary parlance, these random aspects of survival and reproduction are called genetic drift. In large populations, the fluctuations caused by genetic drift are relatively small and tend to be overwhelmed by the systematic hill-climbing effect of natural selection. In small populations, however, these random fluctuations can be more important. Of particular interest here, individuals with lower fitness (off the current peak) may replace those of higher fitness (on the current peak). This process reduces the population's mean fitness, but it sometimes also allows the population to cross a fitness valley and discover another, possibly higher, fitness peak (Whitlock, 1995). This effect of small population size was a central part of Wright's Shifting Balance Theory (Wright, 1932, 1982), in which random genetic drift allows populations to move between fitness peaks.

To see this effect, start by drawing a single fitness peak of moderate height (darkness) and allow a large population to converge on it. Set the mutation effect size to be very small (between 5 and 10), and then draw a second higher (darker) peak that is separated from the first peak by a narrow valley, as shown in Figure 3 A. Notice that this large population stays centered on the first, lower peak because there is selection against genotypes in the low-fitness valley. Now lower the population size to about 10 or 20 individuals and observe how the population becomes much more dynamic, in the sense that its center of mass frequently wanders away from the center of the first peak (Fig. 3 B). The population will occasionally even fall off the peak, so that several individuals can be found in the fitness valley between the two peaks. After some time, the population may move onto the second peak, having crossed the valley that was impassable by the larger population.

Density-dependence and Diversity

In the previous exercises, most or all individuals ended up in one region of the fitness landscape, which means there was very little genetic diversity. But the biological world is incredibly diverse, so we would like to understand how evolution produces and sustains that diversity. There are many factors that affect biological diversity, and in this exercise we will demonstrate one important factor that concerns the nature of interactions among organisms. Density-dependence refers to biological processes for which the rates depend on the density of organisms. For example, in the familiar model of logistic population growth, the per capita
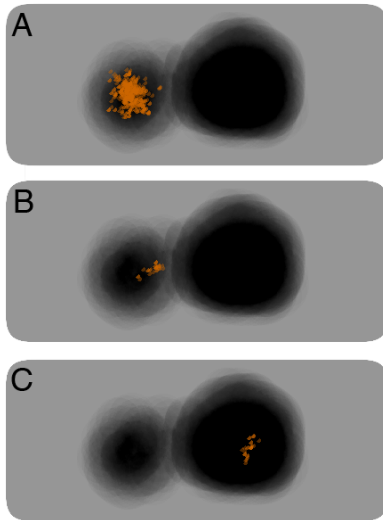
Figure 3: Demonstration of the effect of genetic drift in small populations. Start with a large population that has already converged on a fitness peak. Then add a second, higher peak that is nearby but separated from the first peak by a valley, so that the population does not exhibit the hill-climbing behavior shown in Figure 2. (A) In the large population, new genotypes that are off the peak are quickly replaced by more fit individuals. (B) Now reduce the population size to a very few individuals. New genotypes that are off the peak are not replaced as quickly and some may fall into the basin of attraction of the second peak. (C) After the second peak has been colonized, the population will then typically exhibit the familiar hill-climbing dynamic.

rate of reproduction declines as the population density increases. Density-dependent selection refers to situations in which the fitness of genotypes depends on the number of interactions between individuals. In the case of negative density-dependence, the fitness of an individual declines when it has more interactions with other individuals. (Frequency-dependent selection is a similar concept. Because population density is constant in our application, except when changed by the user, frequency-dependent and density-dependent effects are equivalent.) Negative density-dependent effects often result from increased competition for resources, but they can also result from interactions with predators or parasites whose density increases with that of their prey or hosts. In the context of fitness landscapes, we expect these negative interactions to be more intense among similar genotypes than among those that are dissimilar. In this exercise, we show how that variation in interaction strength promotes diversity by allowing subpopulations to coexist on multiple fitness peaks.

For this exercise, begin by drawing two adjacent fitness peaks of different height. Make sure that the population
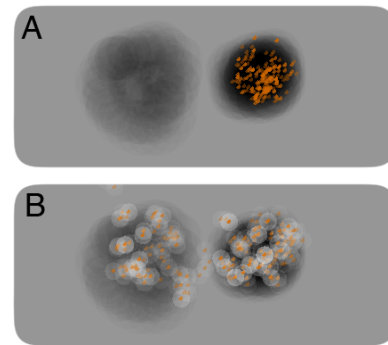

Figure 4: Demonstration of negative density-dependent selection and its effect on diversity. Draw two nearby but distinct fitness peaks of unequal height. Set both population and mutation effect sizes to intermediate or high values, then randomize the population. (A) With density-dependence turned off, the entire population will converge on the higher peak. (B) Now activate the toggle for density-dependence, and observe how the population spreads out and occupies both peaks.

size and mutation effect size are both restored to intermediate or high values (not kept at the small values from the previous exercise). After randomizing the population, the vast majority of individuals will soon occupy only the higher (darker) peak, as seen in Figure 4 A, because individuals on the higher peak produce more offspring than those on the lower peak. Now activate density-dependence and observe that a lighter region surrounds each individual. The lighter color indicates a depression in the fitness landscape relative to the level if that individual were not there. Notice, too, that this effect increases when multiple individuals are in close proximity. Now watch as the population spreads out, first over the current peak and then onto the second peak, as illustrated in Figure 4 B. This shift occurs because the individuals on the first peak depress their own fitness to the point that the second peak becomes the higher one. The two subpopulations - species, perhaps - will then coexist indefinitely.

Survival of the Flattest at High Mutation Rates

Evolution is often described colloquially as survival of the *fittest*. That is, genotypes with high fitness tend to produce more offspring and thereby propel the population up a local peak, as we saw in the first exercise. However, if the peak is very narrow and mutation effects are large, then high-fitness individuals tend to produce offspring that have fallen off the peak and thus have low fitness. In that case, selection may favor genotypes that are *less* fit, in the sense of producing fewer offspring, but more robust because mutations tend to have less harmful effects on their offspring. This scenario has been dubbed "survival of the flattest" because the more

robust types occupy lower but flatter regions of the fitness landscape rather than high but narrow peaks (Wilke et al., 2001). This phenomenon is thought to be important in both computational and biological systems (Wilke et al., 2001; Wilke and Adami, 2003; Beardmore et al., 2011).
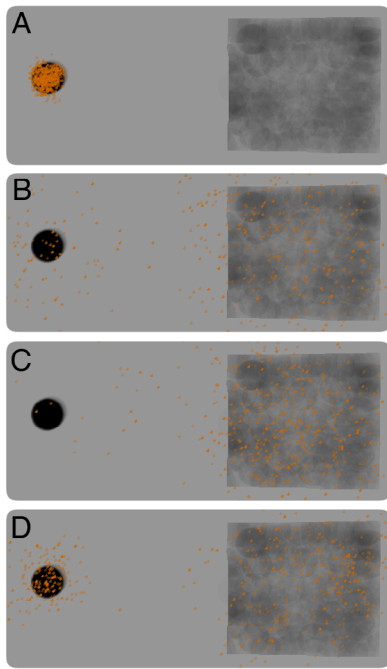

Figure 5: Demonstration of survival of the flattest. (A) Start with a high but narrow fitness peak, and allow the population to converge on it. Set the mutation effect size to a low value. Now draw a second peak that is much lower and broader than the first peak. (B, C) Gradually raise the mutation effect size, and watch as the population moves to the second flatter peak. (D) Now suddenly reduce the mutation effect size to a low value, and the population may move back to the high but narrow fitness peak.

Survival of the flattest is easy to demonstrate, even though the effect is a recent discovery. Figure 5 shows the setup, with a single high but narrow fitness peak, and a much broader but lower peak or plateau. When the mutation effect size is small, the population will remain tightly centered on the high but narrow peak (Figure 5 A). But as you gradually raise the mutation effect size, notice that the population becomes a progressively larger cloud, with many low-fitness offspring born off the peak (Figure 5 B). As you increase the mutation effect size even more, the population abandons the high but narrow peak entirely, provided the two peaks are far enough apart, and spreads out across the lower, flatter peak (Figure 5 C). If the mutation rate is suddenly reduced back to a low level, the population may shift back to the high but narrow peak (Figure 5 D), although this reversal also depends on the distance between the two peaks in relation to other parameters.

Further Explorations

In all of the previous exercises, the mode of reproduction was asexual, which is the default when one begins the application. The interested reader might want to repeat the previous exercises, except with sexual reproduction enabled using the toggle on the display screen. In what cases are the outcomes similar for asexual and sexual reproduction, and when do they differ? We would suggest, in particular, that readers explore the effects of reproductive mode in combination with density-dependent effects. We observed before that density-dependent interactions induced asexual populations to diversify and thereby occupy multiple peaks, as though the subpopulations had split into distinct species. With sexual reproduction, however, intermediate forms (hybrids) are continually generated. To explore the consequences, the reader can switch back and forth between asexual and sexual modes of reproduction, add and remove peaks, and so on.

Conclusions

We built an interactive visualization system that allows users to create fitness landscapes by finger-painting them on a blank canvas. By doing so, Wright's largely metaphorical fitness landscape becomes a playground where one can hone intuition for more formal future experimentation and analysis. Our system is effective for building intuition because all of the processes are visual and mechanistic, while the entire process can be watched in real-time. In addition to painting the initial landscape, users can interact with the system by adding or erasing fitness peaks and by changing parameters such as mutation effect size. We outlined several examples that span a range of complexity from educational exercises to actively researched topics.

Acknowledgments

References

Beardmore, R. E., Gudelj, I., Lipson, D. A., and Hurst, L. D. (2011). Metabolic trade-offs and the maintenance of the fittest and the flattest. *Nature*, 472:342–346.

Gavrilets, S. (2004). *Fitness Landscapes and the Origin of Species*. Princeton University Press, Princeton, N.J.

Kashtan, N., Noor, E., and Alon, U. (2007). Varying environments can speed up evolution. *Proceedings of the National Academy of Sciences*, 104:13711–13716.

Kimura, M. (1983). *The Neutral Theory of Molecular Evolution*. Cambridge University Press, New York, N.Y.

Langdon, W. (1999). Scaling of program fitness spaces. *Evolutionary Computation*, 7:399–428.

Langdon, W. and Poli, R. (1999). Boolean functions fitness spaces. *Genetic Programming*, pages 651–652.

Orr, H. (1998). The population genetics of adaptation: the distribution of factors fixed during adaptive evolution. *Evolution*, pages 935–949.

Pigliucci, M. and Kaplan, J. M. (2006). *Making Sense of Evolution: The Conceptual Foundations of Evolutionary Biology*. University of Chicago Press, Chicago.

Provine, W. (1989). *Sewall Wright and Evolutionary Biology*. University of Chicago Press.

Simpson, G. G. (1953). *The Major Features of Evolution*. Columbia University Press, New York, N.Y.

Whitlock, M. C. (1995). Variance-induced peak shifts. *Evolution*, 49:252–259.

Whitlock, M. C., Phillips, P. C., Moore, F. B.-G., and Tonsor, S. J. (1995). Multiple fitness peaks and epistasis. *Annual Review of Ecology and Systematics*, 26:601–629.

Wilke, C. O. and Adami, C. (2003). Evolution of mutational robustness. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 522:3–11.

Wilke, C. O., Wang, J. L., Ofria, C., Lenski, R. E., and Adami, C. (2001). Evolution of digital organisms at high mutation rates leads to survival of the flattest. *Nature*, 412:331–333.

Wright, S. (1932). The roles of mutation, inbreeding, crossbreeding and selection in evolution. In *Proceedings of the 6th International Congress of Genetics*, volume 1, pages 356–366.

Wright, S. (1982). The shifting balance theory and macroevolution. *Annual Review of Genetics*, 16:1–20.

Wright, S. (1988). Surfaces of selective value revisited. *The American Naturalist*, 131:115–123.