REPLAYING LIFE'S VIRTUAL TAPE: EXAMINING THE ROLE OF HISTORY IN EXPERIMENTS WITH DIGITAL ORGANISMS

By

Jason Nyerere Bundy

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Integrative Biology—Doctor of Philosophy
Ecology, Evolutionary Biology and Behavior—Dual Major

2021

ABSTRACT

REPLAYING LIFE'S VIRTUAL TAPE: EXAMINING THE ROLE OF HISTORY IN EXPERIMENTS WITH DIGITAL ORGANISMS

By

Jason Nyerere Bundy

Evolution is a complex process with a simple recipe. Evolutionary change involves three essential "ingredients" interacting over many generations: adaptation (selection), chance (random variation), and history (inheritance). In 1989's *Wonderful Life*, the late paleontologist Stephen Jay Gould advocated for the importance of historical contingency—the way unique events throughout history influence future possibilities—using a clever thought experiment of "replaying life's tape". But not everyone was convinced. Some believed that chance was the primary driver of evolutionary change, while others insisted that natural selection was the most powerful influence. Since then, "replaying life's tape" has become a core method in experimental evolution for measuring the relative contributions of adaptation, chance, and history. In this dissertation, I focus on the effects associated with history in evolving populations of digital organisms—computer programs that self-replicate, mutate, compete, and evolve in virtual environments. In Chapter 1, I discuss the philosophical significance of Gould's thought experiment and its influence on experimental methods. I argue that his thought experiment was a challenge to anthropocentric reasoning about natural history that is still popular, particularly outside of the scientific community. In this regard, it was his way of advocating for a "radical" view of evolution. In Chapter 2—Richard Lenski, Charles Ofria, and I describe a two-phase, virtual, "long-term" evolution experiment with digital organisms using the Avida software. In Phase I, we evolved 10 replicate populations, in parallel, from a single genotype for around 65,000 generations. This part of the experiment is similar to the design of Lenski's *E. coli* Long-term Evolution Experiment (LTEE).

We isolated the dominant genotype from each population around 3,000 generations (shallow history) into Phase I and then again at the end of Phase I (deep history). In Phase II, we evolved 10 populations from each of the genotypes we isolated from Phase I in two new environments, one similar and one dissimilar to the old environment used for Phase I. Following Phase II, we estimated the contributions of adaptation, chance, and history to the evolution of fitness and genome length in each new environment. This unique experimental design allowed us to see how the contributions of adaptation, chance, and history changed as we extended the depth of history from Phase I. We were also able to determine whether the results depended on the extent of environmental change (similar or dissimilar new environment). In Chapter 3, we report an extended analysis of the experiment from the previous chapter to further examine how extensive adaptation to the Phase I environment shaped the evolution of replicates during Phase II. We show how the form of pleiotropy (antagonistic or synergistic) between the old (Phase I) and new (Phase II) habitats was influenced by the depth of history from Phase I (shallow or deep) and the extent of environmental change (similar or dissimilar new environment). In the final chapter Zachary Blount, Richard Lenski, and I describe an exercise we developed using the educational version of Avida (Avida-ED). The exercise features a two-phase, “replaying life’s tape” activity. Students are able to explore how the unique history of founders that we pre-evolved during Phase I influences the acquisition of new functions by descendent populations during Phase II, which the students perform during the activity.

I dedicate this dissertation to the memory of my friend Cody, a.k.a. “Wicked Rebel”, and to his mother, “Sacred Mom”.

ACKNOWLEDGMENTS

It is easier for me to write about my science than to write about the people in my life. But I need to express my gratitude for the people who made writing this dissertation possible. My career is the result of their investment in my future.

It is hard to find the words to explain the influence my professors have had on my life. Honestly, I avoid the subject because it makes me uncomfortably emotional. But my teachers and advisers are undoubtedly the cornerstone of my motivation to be a tenured faculty member at a major research institution. They are the people who deserve the most credit for what I have been able to achieve academically.

I have had the great privilege of completing my doctoral program in the lab of Richard Lenski. Truthfully, I can't write about Rich without crying. I feel completely incapable of describing the tremendous influence he has had on me. But I have to say that he has been there for me through some very difficult times, personally and professionally. He stuck with me through major changes in my research agenda. He has been supportive through challenges with my physical and mental health. He has always encouraged me to be myself and allowed me to be creative. He consistently treats me like a family member. I completely adore his wife, Madeleine. I am so thankful for the kindness they have shown me through the years. My favorite memory from graduate school is being in the Lenski house for Thanksgiving dinner. I am excited to graduate, but *clearly* in no rush to leave the Lenski Lab.

I have an amazing relationship with the other members of my committee. Charles Ofria and Louise Mead (along with Rich) helped recruit me to MSU. They have been there for me ever since. We have also spent quality time and holidays together. These are great memories I will always keep with me. I am so thankful for their dedication and tremendous expertise. Arend Hintze

taught me the essential programming skills I needed to make my research possible. I had no clue what I was doing when I got here. I've certainly made some missteps along the way. But I put together a tremendous committee. That is one thing I got right the first time.

I would also like to acknowledge my mentors from The Pennsylvania State University. I went to Penn State after graduating from a small Catholic school with 50 students in my class. I was overwhelmed at a campus with over 40,000 undergraduates. Maintaining a close relationship with my professors made one of the largest schools in the country feel small and intimate. My teachers made me feel welcome in the academic community. Diana Gruendler was my Honors English teacher. She helped me find my voice and encouraged me to write about topics that were important to me. Sociologists Sam Richards and his wife, Laurie Mulvey, treated me like the son they never had. I trained facilitators for their Race Relations Project and was a teaching assistant for their popular Sociology of Race class. They were there for me in the early stages of my struggles with mental health, and my family is grateful for the support they showed me. Mark Shriver is a Professor of Anthropology who mentored me and encouraged me to share my research at the Evolution conference that led to me coming to Michigan State. I used to record my tv show in his basement, and he taught me how to use a PCR machine! Richard Doyle is an English professor whose Rhetoric of Science class fueled my love of scientific storytelling and reading books by important scholars like Charles Darwin. But he was also my best friend. I remember meditating on the floor of his office with other students when it was too cold to go outside and playing in his garden along with his children. We stayed close during the years I had to pause my formal education. He advocated for my reinstatement to the scholarship I lost due to my challenges with bipolar disorder. He is also the person who encouraged me to study evolutionary biology. I received my master's degree in biological anthropology in the lab of David Puts. David originally

encouraged me to join his lab a few semesters earlier as an undergraduate research assistant after I returned to Penn State. He helped me develop an independent research agenda and encouraged me to turn my passion for research into a career.

I would also like to acknowledge the other members of the Lenski Lab. I have learned a great deal from them and have wonderful memories with every undergraduate researcher, graduate student, postdoctoral researcher, and guest. I would especially like to acknowledge our outstanding lab managers, Neerja Hajela and Devin Lake. I would like to thank Rohan Maddamsetti for spending so much time with me when I arrived in East Lansing. I want to thank Alita Burmeister for giving me *E. coli* homework, Mike Wiser for teaching me how to use a pipette, and Zack Blount for our lengthy email exchanges about contingency.

I want to say a special thank you to Nkrumah Grant and Kyle Card, my pact brothers. Selfishly, when I met Nkrumah and Kyle, I thought Rich recruited them into the lab just for me! I related to both tremendous scholars in unique ways, and it was great to feel like I was part of a cohort in our lab.

I have also had the great fortune of working at the BEACON Center for the Study of Evolution in Action for the last eight years. BEACON has meant everything to me. BEACON has been a consistent source of personal, professional, and financial support. I would especially like to thank Erik Goodman, Rob Pennock, Chris Adami, Danielle Whittaker, and our former Diversity Director, Judi Brown Clarke for their support. I am also grateful for the patience, guidance, and companionship of my dear friend, Tim Schmidt.

I would like to thank the support staff from the Department of Integrative Biology, the Institute for Cyber-Enabled Research, the Digital Evolution Lab, and the Ecology, Evolutionary

Biology, and Behavior Program. I would not have made it through graduate school without the help of Connie James and Lisa Craft.

I would like to say a special thank you to the communities that have supported me throughout my time in graduate school. I want to thank the members of the Sacred Death Motorcycle Club for their unconditional love and support. I would also like to thank members of the Commons Church and its college ministry, the Salt Company. I have always wanted to be the drummer for an active campus ministry. Thank you so much for making that ambition a reality. I loved every second of it!

I want to thank my family. My grandparents, Moses and Edythe Priester, have been very involved in my life. I am thankful for the stable foundation and noble example they have provided our family. My parents, Wanda Priester and Al Bundy, have made countless sacrifices to provide me with every opportunity to succeed in life. My passion for higher education was fueled by an early love of spending time on college campuses. That love began with spending time with my father and his students at Seton Hall University and sharing homework time with my mom as she earned her master's degree at the University of Scranton. My family has always had way too much confidence in my abilities! I remember when my father left his position at Seton Hall before my senior year in high school. I was stunned. Why would he do that when I so close to being able to attend Seton Hall for free!? When I asked him, he said, "You're going to get a scholarship to go to college." He was right.

Thanks for believing in me. I love you all!

TABLE OF CONTENTS

LIST OF TABLES xii

LIST OF FIGURES xiii

CHAPTER 1: WONDERFUL LIE: "REPLAYING LIFE'S TAPE" AND EVOLUTIONARY FATALISM 1
The Good News According to Gould: The Great Gedankenexperiment 2
The Purpose Driven History of Life 3
Replaying a Tape from Gould's Wonderful Life 10
Contingency Before Gould: From Epigenesis to Evolutionary Fatalism 13
Toward a New Iconography 19
The Experimental Era: Replaying Life's Tape in the Laboratory 23
REFERENCES 27

CHAPTER 2: HOW THE FOOTPRINT OF HISTORY SHAPES THE EVOLUTION OF DIGITAL ORGANISMS 34
Abstract 35
Introduction 36
Experimental Design 39
Results 42
Evolution in Phase I 42
Phase II: Evolution in the Overlapping environment after a shallow Phase I history 43
Trajectories for fitness and genome length 43
Contributions of adaptation, chance, and shallow history 44
Phase II: Evolution in the Orthogonal environment after a shallow Phase I history 46
Trajectories for fitness and genome length 47
Contributions of adaptation, chance, and shallow history 47
Effects of deepening the footprint of Phase I history on evolution during Phase II 48
Effect of deepening the footprint of history on genome length 49
Effect of deepening the footprint of history on fitness 52
Confidence intervals and confident inferences 55
Discussion 57
Overview 57
Future directions 60
Concluding thoughts 63
Materials and Methods 63
Avida system and experiments 63
Analysis 66
Acknowledgments 67
APPENDIX 68
REFERENCES 77

CHAPTER 3: HABITAT NOVELTY AND HISTORY SHAPE TRADE-OFFS AND TRADE-UPS IN DIGITAL ORGANISMS 83
Abstract 84
Introduction 85
Results 92
Phase I: Evolving founders with shallow and deep history in the old environment. 92
Phase II: Habitat novelty shapes different responses between founders with shallow and deep history. 94
Fitness of founders, gain-of-function genotypes, and final dominant genotypes in the old and new environments. 95
Fitness of the final dominant genotype and its founder in the old and new environments 99
Correlated fitness changes in the old and new environments. 103
Contingency and chance both contribute to the among-population variation in pleiotropic effects. 105
Discussion 108
Future Directions 110
Materials and Methods 115
Avida system and original experiment 115
Analysis 116
APPENDIX 120
REFERENCES 125

CHAPTER 4: EXPLORING THE EFFECTS OF HISTORY ON EVOLUTION: AN AVIDA-ED EXERCISE 133
Abstract 134
Introduction 135
Background 136
Developing the Exercise 137
APPENDIX 141
REFERENCES 156

LIST OF TABLES

Table 3.1. The proportion of lineages in the four treatments that exhibited cumulative fitness losses in the ancestral environment, indicative of antagonistic pleiotropy 102

LIST OF FIGURES

Figure 1.1. “The Great Chain of Being” by Didacus Valades from *Rhetorica Christiana* 6

Figure 1.2. The Road to Homo Sapiens, a.k.a. “The March of Progress” 7

Figure 1.3. The “cone of increasing diversity” vs “decimation and diversification” 8

Figure 1.4. A modern “tree of life” representation..... 20

Figure 1.5. The “tree of life” based on genome context networks 21

Figure 1.6. A modern depiction of hominin evolution..... 22

Figure 2.1. Schematic illustration of the contributions of adaptation, chance, and history..... 38

Figure 2.2. Schematic illustration of the experimental design..... 41

Figure 2.3. Population averages for fitness and genome length during Phase I 43

Figure 2.4. Population averages for fitness and genome length during Phase II in the Overlapping environment for populations with a shallow history..... 44

Figure 2.5. End-of-experiment effects of adaptation, chance, and history in the Overlapping environment with shallow history..... 45

Figure 2.6. Trajectories of adaptation, chance, and history in the Overlapping environment with shallow history 46

Figure 2.7. Impact of the footprint of history on the contributions of adaptation, chance, and history to genome length is similar in the two Phase II environments 51

Figure 2.8. Impact of the footprint of history on the contributions of adaptation, chance, and history to fitness differs between the two Phase II environments..... 53

Figure 2.9. Trajectories for the contributions of adaptation, chance, and history to fitness in the new Phase II environment..... 55

Figure 2.10. Population averages for fitness and genome length during Phase II in the Orthogonal environment for populations with a shallow history..... 69

Figure 2.11. End-of-experiment effects of adaptation, chance, and history in the Orthogonal environment with shallow history..... 70

Figure 2.12. Trajectories of adaptation, chance, and history in the Orthogonal environment with shallow history 71

Figure 2.13. Population averages for fitness during Phase II with intermediate history 72

Figure 2.14. Population averages for genome length during Phase II with intermediate history. 73

Figure 2.15. Population averages for fitness during Phase II with deep history 74

Figure 2.16. Population averages for genome length during Phase II with deep history 75

Figure 2.17. Trajectories for the contributions of adaptation, chance, and history to genome length during Phase II 76

Figure 3.1. Schematic illustration of the experimental design..... 87

Figure 3.2. Fitness trajectories during Phase I 88

Figure 3.3. Impact of deep history on adaptation depends on extent of environmental change... 90

Figure 3.4. Average fitness in the Phase I environment of the Phase II founders with shallow or deep history in that environment..... 93

Figure 3.5. Average number of functions performed by the Phase II founders with shallow or deep history 94

Figure 3.6. Average fitness of Phase II founders with shallow or deep history measured in the new Overlapping and Orthogonal environments..... 95

Figure 3.7. Fitness during Phase II for the founders, intermediate gain-of-function genotypes, and final dominant genotypes that evolved in the Overlapping environment with shallow history ... 97

Figure 3.8. Fitness of each population's founder and final dominant genotype in the old and new environments..... 100

Figure 3.9. Fitness of each final dominant genotype relative to its founder in the old and new environments..... 101

Figure 3.10. Fitness differentials in the old and new environments between each final dominant genotype and its founder..... 104

Figure 3.11. Correlation between fitness changes in the old and new environments during Phase II 105

Figure 3.12. History and chance contribute to variation in correlated responses..... 107

Figure 3.13. Fitness during Phase II for the founders, intermediate gain-of-function genotypes, and final dominant genotypes that evolved in the Overlapping environment with deep history 121

Figure 3.14. Fitness during Phase II for the founders, gain-of-function intermediate genotypes, and final dominant genotypes that evolved in the Orthogonal environment with shallow history... 122

Figure 3.15. Fitness during Phase II for the founders, intermediate gain-of-function genotypes, and final dominant genotypes that evolved in the Orthogonal environment with deep history 123

Figure 3.16. History and chance contribute to variation in the proportion of antagonistic gain-of-function mutations 124

CHAPTER 1: WONDERFUL LIE: “REPLAYING LIFE’S TAPE” AND EVOLUTIONARY FATALISM

The Good News According to Gould: The Great Gedankenexperiment

In 1989, paleontologist Stephen Jay Gould advocated for the importance of *historical contingency* in evolution by describing a now-famous thought experiment in his book *Wonderful Life*:

> We live … in a world of good news and bad news. The good news is that we can specify an experiment to decide between the conventional and the radical interpretations of extinction, thereby settling the most important question we can ask about the history of life. The bad news is that we can't possibly perform the experiment.
>
> I call this experiment "replaying life's tape." You press the rewind button and, making sure you thoroughly erase everything that actually happened, go back to any time and place in the past–say, to the seas of the Burgess Shale. Then let the tape run again and see if the repetition looks at all like the original. If each replay strongly resembles life's actual pathway, then we must conclude that what really happened pretty much had to occur. But suppose that the experimental versions all yield sensible results strikingly different from the actual history of life? What could we then say about the predictability of self-conscious intelligence? or of mammals? or of vertebrates? or of life on land? or simply of multicellular persistence for 600 million difficult years? [1]

My dissertation is focused on "replaying life's tape" with digital organisms to measure historical effects in evolution. In this chapter, I begin with a discussion of the philosophical significance of Gould's great *gedankenexperiment* [German: thought experiment]. That is followed by an overview of its influence on the techniques used in the remainder of this dissertation.

First, I discuss what Gould referred to as the "conventional*"* view. Then I contrast it with what he called the "radical*"* interpretation. I describe why the distinction between these perspectives is elusive in the history of science.

In *Wonderful Life,* Gould opposed to the anthropocentric "iconography of progress" that was part of the "conventional" view. He identified core presuppositions of this iconography. He discussed the challenges to these presuppositions posed by his understanding of evolutionary theory. I will also describe the revised iconography of more recent visualizations.

Today, the idea of "replaying life's tape" is central to the research program that investigates historical contingency. Some of the approaches now widely used resemble Gould's thought experiment. As a result, there is a growing body of evidence beyond the pages of *Wonderful Life*. Thus, Gould's thought experiment is noteworthy in the history of both evolutionary philosophy and method.

The Purpose Driven History of Life

> "All the evidence available in the biological sciences supports the core proposition … That the cosmos is a specially designed whole with life and mankind as its fundamental goal and purpose, a whole in which all facets of reality have their meaning and explanation in this central fact."- Michael Denton [as quoted in Rick Warren's *The Purpose Driven Life*] [2]

The philosophical significance of Gould's thought experiment goes beyond demonstrating alternative facts that might result from a repetition of natural history. Gould's zeal for historical contingency and the motivation for his thought experiment were, in part, to illustrate "radical" aspects of his reinterpretation of Charles Darwin's theory. Before cueing up Gould's explication of the revolution, I should establish the old order that it claims to upend.

Philosopher Daniel Dennett used the proceeding quotation as part of his TED talk responding to pastor Rick Warren [3]. Warren featured this quote from intelligent-design proponent Michael Denton in his book. Dennett argued forcefully against the use of intelligent design as an explanation for human "purpose." However, I was particularly drawn to the phrase "all the evidence available in the biological sciences supports the core proposition." Following Dennett's lead, I could dismiss Denton's language as an effort to push the creationist agenda by using hyperbolic rhetoric to subsume the facts of natural history. But I would miss an opportunity to address this as a quasi-scientific claim that supports an anthropocentric interpretation of natural history. This anthropogenesis-centered perspective is similar in that respect to the view Gould sought to address when critiquing what he called the "conventional" interpretation. Recognizing this concordance might be productive in discourse with those who genuinely accept the *fact* of evolution, but who wrestle with the implications of evolutionary *theory* [4]. Dennett may be correct about Warren and Denton's creationist leanings. But perhaps one does not need to be a creationist to believe the available scientific evidence supports Denton's conclusion about human origins.

Gould's thought experiment goes beyond providing a framework to contrast these "radical" and "conventional" interpretations of extinction. His explanation of the thought experiment's implications can stimulate a discussion of the "iconography of progress" that is still crucial to how many people see evolution.

The history of science can help us understand the origin of Denton's claim about the available evidence. To Gould, the "conventional view" is recognizable by three related presuppositions. They can produce a tendency to view evolution as a predetermined process. Paleontologist George Gaylord Simpson called this tendency "evolutionary fatalism."

The first presupposition is progress. This is the idea that primitive, implicitly simple organisms like bacteria are "less evolved" or in some way inferior to later—"complex", "sophisticated", and "highly evolved"—creatures like *Homo sapiens*. This intuition makes sense of our recent origin by mapping an anthropocentric "natural" hierarchy onto the timeline inferred from old fossils, an older Earth, and an incomprehensibly old universe. In other words, we may have taken a long time to arrive, but our emergence was the theme all along.

The second presupposition is directionality. Progress occurs along one of a limited number of paths toward a particular end. You may not know exactly where the train is headed when it leaves the station. But you are on track, headed inexorably toward your destination.

The third presupposition is predictability. This does not assume that we can anticipate specific evolutionary events. Rather, the patterns of natural history suggest a peculiar reliability. From that evidence, some infer the inevitability of certain outcomes, including the evolution of human consciousness or something like it. Of course, natural history cannot be repeated. But the structure and regularities of the physical world make the underlying evolutionary process essentially repeatable.

Gould described how these assumptions were crystalized in representations of natural history. He referred to these visual aids and the narratives that support them as the "iconographies of progress" [1]. He described how the imagery, in turn, shapes how we interpret data. I think Gould might have agreed with my tongue-in-cheek description of the biases that result from the "iconographies of progress" as a *wonderful lie*: "The familiar iconographies of evolution are all directed–sometimes crudely, sometimes subtly–toward reinforcing a comfortable view of human inevitability and superiority. The starkest version, the chain of being or ladder of linear progress, has an ancient, pre-evolutionary pedigree" [1].

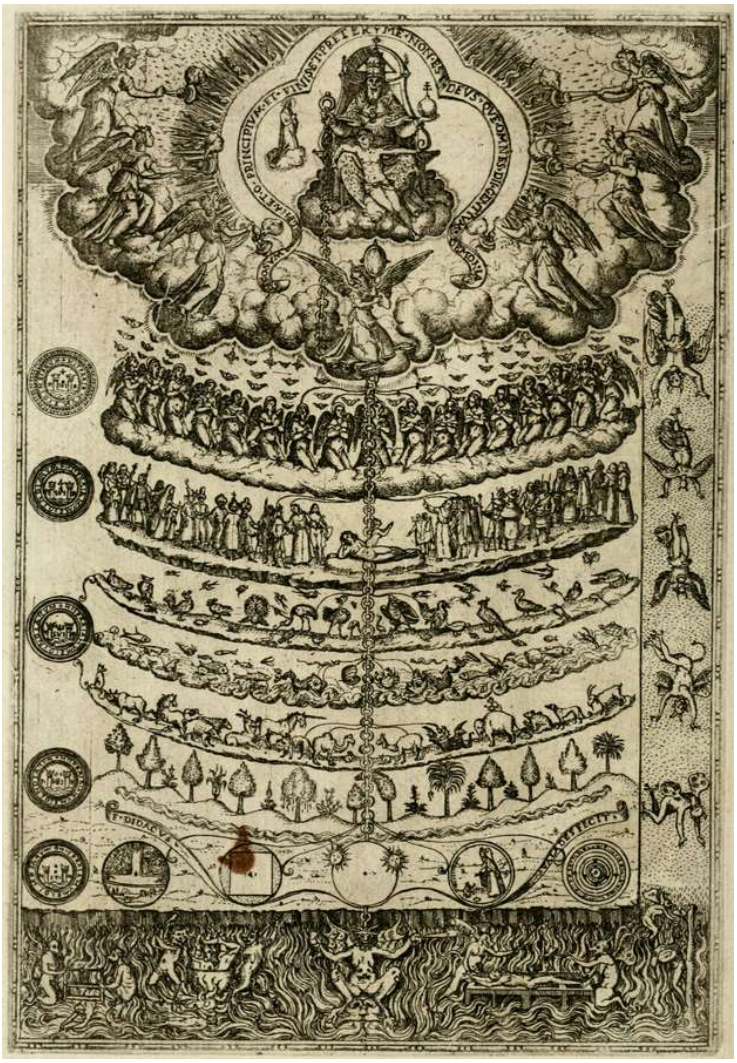


Figure 1.1. "The Great Chain of Being" by Didacus Valades from *Rhetorica Christiana*
The Chain of Being is depicted as a hierarchy with the deity above spiritual beings (angels) [5]. Spiritual beings are higher than humans, who sit atop the levels occupied by animals and plants.

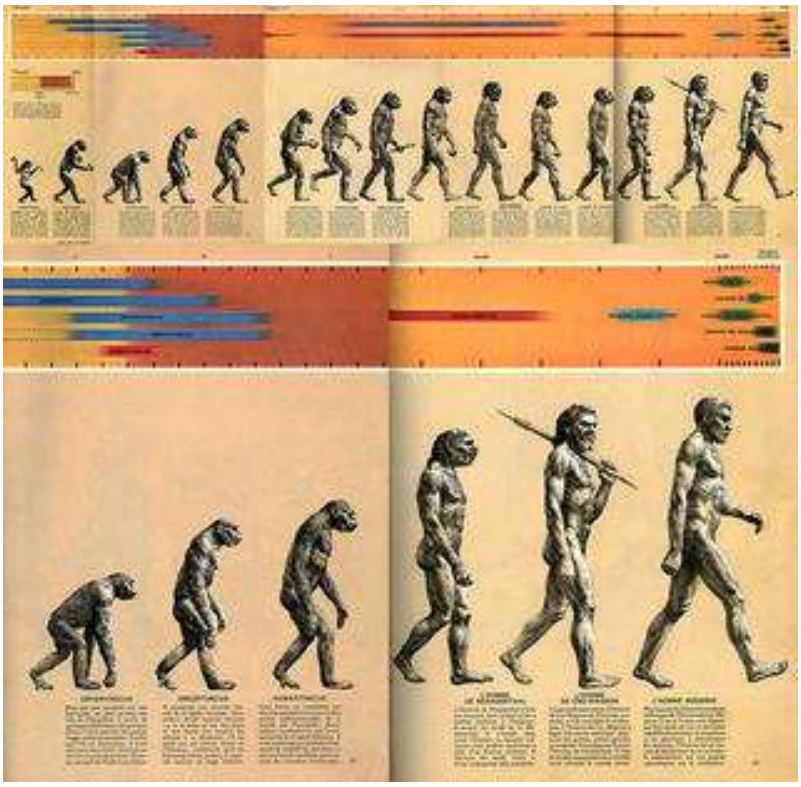


Figure 1.2. The Road to Homo Sapiens, a.k.a. "The March of Progress"
The full version of "The Road to Homo Sapiens" above the folded version from 1965's Early Man [6]. The folded version, in particular, seems to give the appearance of a linear progression leading to modern Homo sapiens. This depiction was criticized for being inconsistent with the branching nature of evolutionary relationships.

Philosopher John Dewey commented on the relationship between these iconographies in a speech celebrating the 100th anniversary of Darwin's birth and the 50th anniversary of the publication of Darwin's *On the Origin of Species* [7]. Dewey noted the "incompleteness of" and "active resistance to" the Darwinian revolution in "the recrudescence of absolutistic philosophies" that "pay lip service to Darwinian evolution by turning on its side the Great Chain of Being, with its hierarchically ordered ranks of essentially defined types ranging from monad to man, and watching it unfold in time" [7,8].

Another key icon in the iconography that Gould described is the "cone of increasing diversity" [1]. He contrasts the cone with his own view of "decimation and diversification" based

on his interpretation of the Cambrian Burgess Shale and the fossil record more generally [1]. Gould observed a wide array of early forms amongst the Burgess Shale fauna (an attribute he called "disparity") [1]. He thought many of these forms were on lineages distinct from modern evolutionary branches [1]. In Gould's view, most of these branches were wiped out by the arbitrary nature of extinction, not because they were inferior [1].

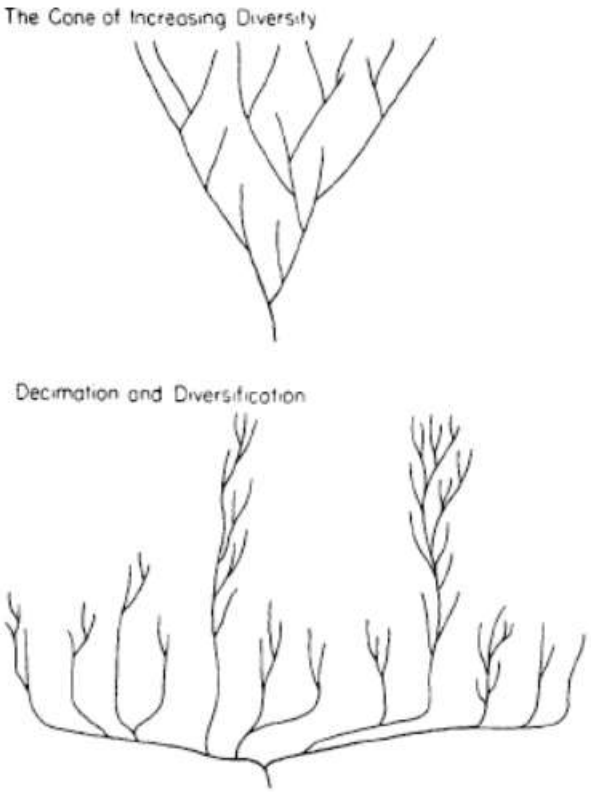


Figure 1.3. The "cone of increasing diversity" vs "decimation and diversification"
Gould's depiction of the "conventional" view as a "cone of increasing diversity" is contrasted with his own view of "decimation and diversification" based on his analysis of Burgess Shale fauna [1]. Gould's view is distinguished by a much broader array of early forms, which he called "disparity". Many are exclusive to lineages that were annihilated by extinction.

Gould was clear about the "rose-colored glasses" we don if we are not mindful of our assumptions.

> The fatuous idea of a single order amidst the multifarious diversity of modern life flows from our conventional iconographies and the prejudices that nurture them– the ladder of life and the cone of increasing diversity. By the ladder, horseshoe crabs are judged as simple; by the cone, they are deemed old. And one implies the other under the grand conflation … down the ladder also means old, while low on the cone denotes simple.
>
> I don't think that any particular secret, mystery, or inordinate subtlety underlies the reasons for our allegiance to these false iconographies of ladder and cone. They are adopted because they nurture our hopes for a universe of intrinsic meaning defined in our terms. [1]

Now that the core tenants and iconographies of what Gould called the "conventional" interpretation have been established, the stage is set to consider a "radical" alternative. I will share an insight I encountered while drafting this manuscript. When I revisited the quotation from the beginning of this section, I imagined a seemingly innocuous but non-synonymous mutation—one that gives us "cone" from "core". This little mutation reveals the similarity between the "conventional" reasoning that Gould sought to overturn with his thought experiment and Michael Denton's anthropocentric statement featured in *The Purpose Driven Life*.

> "All the evidence available in the biological sciences supports the *cone* proposition … That the cosmos is a specially designed whole with life and mankind as its fundamental goal and purpose, a whole in which all facets of reality have their meaning and explanation in this central fact."

Replaying a Tape from Gould's Wonderful Life

In 2000, as part of the Distinguished Lecture Series titled "Evolution Revolution", Gould spoke about what he considered revolutionary aspects of Darwin's theory.

> One of the most interesting aspects of Darwin's life, and the most revealing, is that he developed the theory of natural selection in pretty much its complete form … in 1838 and he didn't publish until 1859. That's one of the long-standing puzzles in Darwinian scholarship. Why does he delay 21 years? The old, heroic explanation that he was just collecting his facts, he didn't want to jump the gun, that simply can't be right … it must have been fear and not just mere collection. I don't think fear of exposing evolution, it's not that heretical a view, but of exposing the radical content of his own take on evolution. [9]

Gould compared the Darwinian revolution to the Copernican [9]. He said there is greater cultural resistance to Darwin's theory because the implications are much bigger. He clarified that the Copernican revolution, with its resulting heliocentrism, "is about real estate [9]" whereas the Darwinian revolution "is about essence" [9]. Gould summarized three aspects of Darwin's "radical content."

First, Gould discussed "the radical nature of the theory itself" [9]. Darwin's theory contradicted the popular teleology of evolutionary progress (orthogenesis). Darwin described an entirely natural mechanism, namely natural selection, which produces exclusively local adaptation [9]. Gould suggested that Darwin's initial aversion to the term *evolution* was related to the association of the word with notions of progress in the vernacular [9]. Under Gould's revision of Darwinian thought, organisms are not more or less "evolved". "Highly evolved" implies a hierarchical arrangement of life. To Gould, Darwin's theory demands a non-hierarchical interpretation of natural history.

Populations become suited to a particular environment and way of life over many generations. We refer to this suitability as fitness. The word *fitness* illustrates a tension between Darwinian theory and what Gould considered "conventional" wisdom. Darwinian fitness—the ability to live and reproduce in a particular environment—maps nicely unto the common definition: "the quality of being suitable to fulfil a particular role or task." The English sociologist Herbert Spencer, a proponent of universal progress, popularized the phrase "survival of the fittest" and over time Darwin embraced the use of *evolution (*rather than "descent with modification") to describe his theory [9]. It is easy to misconstrue Darwinian fitness in a way that evokes physical fitness and associated traits like strength, speed, and stamina. But Gould invites us to "take the cold bath": "The parasite that becomes so morphologically simplified that it's just a bag of reproductive tissues in the interior of its host is no less well-adapted than complex 'us' are to our exteriors and we probably each have a reasonably equal chance of extended geological persistence. That's what the theory's about" [9].

The second radical aspect of Darwin's theory, according to Gould, is "what it says about variation" [9]. Darwin's theory predates our modern understanding of DNA and population genetics. Nevertheless, Darwin was able to draw a connection between variation within a population and differences between species. Importantly, he thought the source of this variation was undirected. Variation is generated by chance, without regard to its utility to the individual. Gould emphasizes the unorthodox nature of Darwin's view of variation as follows.

> In Platonic traditions variation is accident, which literally means falling away from the mean (or the essence) of the thing. Variation is traditionally treated as a set of odd accidents of individual, imperfect incarnations of an inherent essence … And therefore, the essence is central and the variation is mere accident. Darwin

> completely inverts that! In Darwinian thought the variation is the essence. What is the essence of the species, but the spectrum of its variation which is available for conversion into evolutionary change? … That's a profound change in perception and it's … of great practical importance. [9]

Gould's exposition reaches a thematic climax when discussing the third radical aspect of Darwin's theory, "which has to deal with historicity and contingency" [9]. Historical patterns may seem intelligible in hindsight, but they are inherently unpredictable in advance because they are not deterministic [9]. Championing this contingent view of natural history is central to Gould's legacy. It is a challenge to inevitability. Gould revels in representing Darwin's dissident doctrine:

> Under most what I'm calling more user-friendly notions of evolution you may still think … you can construe it as a predictable sequence. Maybe every little detail isn't predictable. There's a broad kind of progress. That is inherent … to this day, most people who exclaim a belief in evolution, who even consider themselves Darwinian still see it that way … The common, the vernacular understanding of evolution is that it's a process that means increasing complexity. If you can't exactly be sure that *Homo sapiens* would have emerged, at least something like it is very likely, because that's what evolution means. But Darwinian evolution isn't that! One of the fascinating things that flows from it because it is only about local adaptation and local adaptation is to immediate environments, immediate environments are changing essentially on random vectors through geological time. So that even if you have the determinism of local adaptation for any moment, since it's a determinism of local adaptation to fluctuating environments that are essentially stochastic through time, it's going to give you a stochastic pattern for

> the history of life. And therefor you're stuck in a true Darwinian world … You're going to be stuck with a view of unpredictable history … That's a scary notion, that actual history is unpredictably contingent.

Contingency Before Gould: From Epigenesis to Evolutionary Fatalism

Gould's fascination with and emphasis on the vagaries of history had many precedents. Discussions of contingency can be traced from the works of theologians and natural philosophers of antiquity to pre-Darwinian evolutionists [10–28]. Philosopher David Depew argues for "a more complicated history of biology" [7]. He explains that the Platonic tradition is not as "essentialist" as it is usually portrayed. Plato's student Aristotle understood that species are lineages, not types [7]. Depew defines contingency as "the opposite of necessity" [7]. He highlights the contingency in Aristotle's understanding of reproduction: "Your coming-to-be presupposes your father's, his coming-to-be does not presuppose yours" [7,17]. Aristotle continued, "Why … does this coming-to-be seem to constitute a rectilinear sequence?" [17] A rectilinear, end-directed model of evolution is central to the popularized orthogenesis overturned by natural selection [29]. In Aristotle, we therefore see contingency as an element of the "conventional" interpretation. Later, English physician William Harvey labelled the Aristotelian view *epigenesis* in the early modern period before Darwin [7].

Like contingency, discussions of chance in biology have an ancient history. Depew traces this history back to the Greek pre-Socratic philosopher Empedocles, Greek philosopher-sage Epicurus, and Roman scientist-poet Lucretius [7]. For them, according to Depew, "chance, adds to contingency … an element of indeterminacy that, whether it is located in the nature of things or only in our ignorance, defies explanation, especially explanation in terms of purposes" [7]. This view seems well-aligned with Gould's indeterminate reinterpretation of Darwinian theory.

However, many other thinkers of antiquity supported a "functional and goal-oriented" view of organisms [7]. They denigrated proponents of chance for supporting the kind of "species mutability" that would later be associated with Darwin [7]. Species mutability was also an unpopular idea with natural theologians [7]. By the early nineteenth century, the pre-Darwinian transmutationists were accused of being proponents of chance and detractors of purposiveness in biology [7]. However, Darwin's use of "chance" was more precise and distinct from that of the ancients [7]. In particular, Darwin's theory was more nuanced and stipulated only that the cause of each variation is independent of its utility to the organism.

The idea of "species fixism" that Gould might have associated with the "essence" of pre-Darwinian thought is a somewhat more recent notion [7]. Depew associates the emergence of "typological essentialism" with "the superimposition of the logical category of species onto biological classification and closely related efforts to involve God in fixing species boundaries" [7]. However, Darwin's theory was not merely natural selection plus "chance" because, as Depew wrote, "In Darwinism, contingency and chance combine with environmental determinism to evolve natural purposiveness" [7].

In the 130 years between Darwin's *Origin of Species* (1859) and Gould's *Wonderful Life* (1989), an indeterminant view of natural history, with organisms' finding a variety of solutions to environmental challenges, was common among key evolutionary thinkers [30–37]. Fortuity, not provision, was the core of Darwin's understanding of "spontaneous or accidental" variation [37]. Many years later, George Gaylord Simpson reasserted the centrality of evolutionary "opportunism" to combat the problematic social and political ideas he attributed to "evolutionary fatalism" [38].

> Biological justification for the totalitarian development of society has also been sought in the doctrines of evolutionary fatalism. Regardless of such labels as "right", "wrong", "good", or "bad", it is argued, this is the inevitable future. Mankind is going this way just as horsekind was going toward Equus throughout the Tertiary. Opposition is as futile and foolish as if the little Eohippus had said, "I am going to be a dinosaur," instead of, "—a horse." Even aside from the fact that this is another false use of analogy, it has been shown that a fatalistic view of evolution has little scientific support. [39]

According to Simpson, "There are two aspects of opportunism: to seize such diverse opportunities as occur [by chance], and when a single opportunity or need occurs, to meet it with what is available, even if this is not the best possible" [32]. Simpson foreshadowed Gould's thought experiment when he wondered if "Cambrian trilobites would have evolved in the same way under the same conditions in the Ordovician" [33,38]. Like Gould, Simpson took the resigned stance "history does not repeat itself" [33,38].

There was more at stake in Gould's thought experiment than biological fantasy. *Wonderful Life* was published during a time of intense debate over the prevalence of various evolutionary mechanisms in biology. These debates were highly politicized outside of science. Gould's views received a lot of attention from both scientists and non-scientists.

In particular, Gould and others [40–44] emphasized the importance of historical contingency in response and even opposition to the "adaptationist" paradigm among many of his contemporaries [45]. In a famous paper, Gould and Lewontin tried to promote "Darwin's pluralistic approach" [46].

> Organisms must be analysed as integrated wholes, with *Baupläne* [German: body plans] so constrained by phyletic heritage, pathways of development and general architecture that the constraints themselves become more interesting and more important in delimiting pathways of change than the selective force that may mediate change when it occurs. [46]

Around that time, the Japanese population geneticist Motoo Kimura and others were championing the role of nonadaptive evolution through random genetic drift [47–50]. According to Kimura, "the great majority of evolutionary changes at the molecular level … are caused not by Darwinian selection but by random drift of selectively neutral or nearly neutral mutants" [51]. Building on population geneticist Sewall Wright's earlier work [52–54], the drift-induced founder effect was recognized as a stochastic force that could, in theory, promote phenotypic diversity and even speciation [55]. But others used the prevalence of convergence—the evolution of similar structures or behaviors in distantly related lineages—to emphasize the power of adaptation to produce optimal phenotypes [56,57].

In *Wonderful Life,* Gould relied heavily on the work of paleontologists Harry Whittington and his former graduate student Simon Conway Morris on the Burgess Shale. Yet Conway Morris became a major critic of Gould's interpretation. Gould's view was based on his critique of the classification of the Burgess Shale fauna attributed to the deposit's discoverer, paleontologist Charles Doolittle Walcott. Gould clarified that Walcott's "traditional" predispositions did not make him anti-Darwinian [1]. Instead, in Darwin, Walcott found support for a view of evolution consistent with his preconceived notions. According to Gould, "Walcott considered himself a Darwinian, expressing by this stated allegiance his strong conviction that natural selection assured the survival of superior organisms and the progressive improvement of life on a predictable

pathway to consciousness [1].” Gould quoted Walcott’s lecture “Searching for the First Forms of Life” (also referenced in [58]). The quotation embodies what Gould called the key preconceptions of “Walcott’s shoehorn.”

> In the early times the Cephalopoda ruled, later on the Crustacea came to the fore, then probably fishes took the lead, but were speedily outpowered by the Saurians. These Land and Sea Reptiles then prevailed until Mammalia appeared upon the scene, since when it doubtless became a struggle for supremacy until Man was created. Then came the age of Invention; at first of flint and bone implements, of bows and arrows and fish-hooks; then of spears and shields, swords and guns, lucifer matches, railways, electric telegraphs. [1]

Gould used “shoehorn” to describe how Walcott placed all Burgess Shale organisms into existing phyla, even when the evidence for such classifications was flimsy. The primary purpose of *Wonderful Life*, therefore, was to challenge Walcott’s “shoehorned” classification, which Gould attributed to the “conventional” interpretation, in contrast with his own view that emphasized historical contingency. The “shoehorn” is characterized by anthropocentric biases, similar to those in Denton’s reasoning about “all the evidence available in the biological sciences.” The conventional rhetoric clearly evokes the “iconography of progress.”

Conway Morris invokes adaptation with the very structure of the universe to argue that evolution suggests an inevitable journey from past to present. Like Denton, for Conway Morris, this “optimistic” view includes “life as a cosmic imperative” [57].

> If brains can get big independently and provide a neural machine capable of handling a highly complex environment, then perhaps there are other parallels, other convergences that drive some groups towards complexity. Could the story of

> sensory perception be one clue that, given time, evolution will inevitably lead not only to the emergence of such properties as intelligence, but also to other complexities, such as, say, agriculture and culture, that we tend to regard as the prerogative of the human? We may be unique, but paradoxically those properties that define our uniqueness can still be inherent in the evolutionary process. In other words, if we humans had not evolved then something more-or-less identical would have emerged sooner or later. [57]

This is the crux of evolutionary fatalism writ large. The overwhelming power of selection is used to argue that "replaying life's tape" would be an exercise in documenting a progression of forms very much like the history of life as we know it. In the words of Conway Morris, "perhaps we can discern inherent within this framework the inevitable and pre-ordained trajectories of evolution?" [57]

Gould's stance against the "adaptationist programme" amplified the same objections to anthropocentrism raised by predecessors including Darwin and Simpson [38]. Darwin resisted the idea of a creator who guided variation in order to ensure the emergence of man in his own image [59,60]. Similarly, Gould balked at the notion that humanity was anything but "wildly improbable" [1]. Gould found this improbability an "exhilarating" basis for both "freedom and consequent moral responsibility" [1]. He is explicit about what is on the line regarding the "essence of history" [9], as reflected in this passage from *Wonderful Life*:

> Remember what is at stake! Our most precious hope for the history of life, a hope that we would relinquish with greatest reluctance, involves the concepts of progress and predictability. Since the human mind arose so late, and therefore threatens to demand interpretation as an accidental afterthought in a quirky evolutionary play,

> we are incited to dig in our heels all the harder and to postulate that all previous life followed a sensible order implying the eventual rise of consciousness. The greatest threat lies in a history of numerous possibilities, each sensible in itself after the fact, but each utterly unpredictable at the outset—and with only one (or very few) roads leading to anything like our exalted state. [1]

Toward a New Iconography

Today, the "iconographies of progress" are largely avoided in evolution education. Instead, depictions like the "tree of life" shown below are popular. The poster, inspired by British evolutionary biologist Richard Dawkins's *The Ancestor's Tale*, uses illustrative branches "to suggest the effect of extinction on diversity" [61]. The poster is reminiscent of Gould's depiction of "decimation and diversification." The consortium that produced the poster nonetheless acknowledges that this "tree of life" is depicted from a human point of view. They say that "this falsely suggests that humans are the ultimate goal of evolution" [61]. They also speculate, "if that asteroid or comet that hit the earth 66 million years ago had instead missed the earth, there might not be a dominant, tool-using, space-faring species on earth. Or if one evolved, it might be a dinosaur, not a mammal" [61].

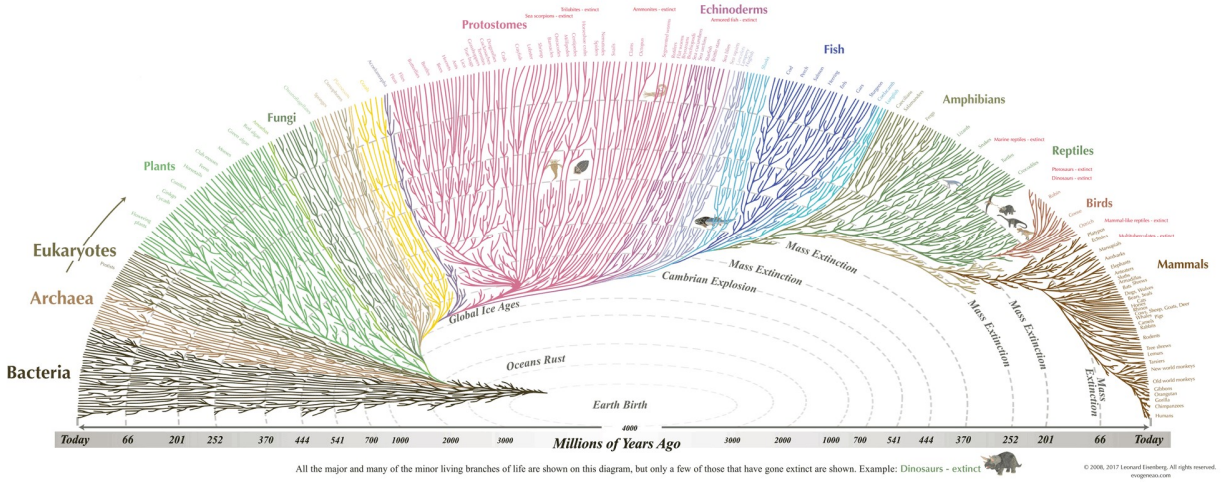


Figure 1.4. A modern "tree of life" representation
This image is from an educational poster used to illustrate the common ancestry between all past and present living organisms [61].

Other modern depictions of the "tree of life" are strictly data-driven and noticeably less anthropocentric. The one shown below is based on genomic analysis of fully sequenced genomes [62].

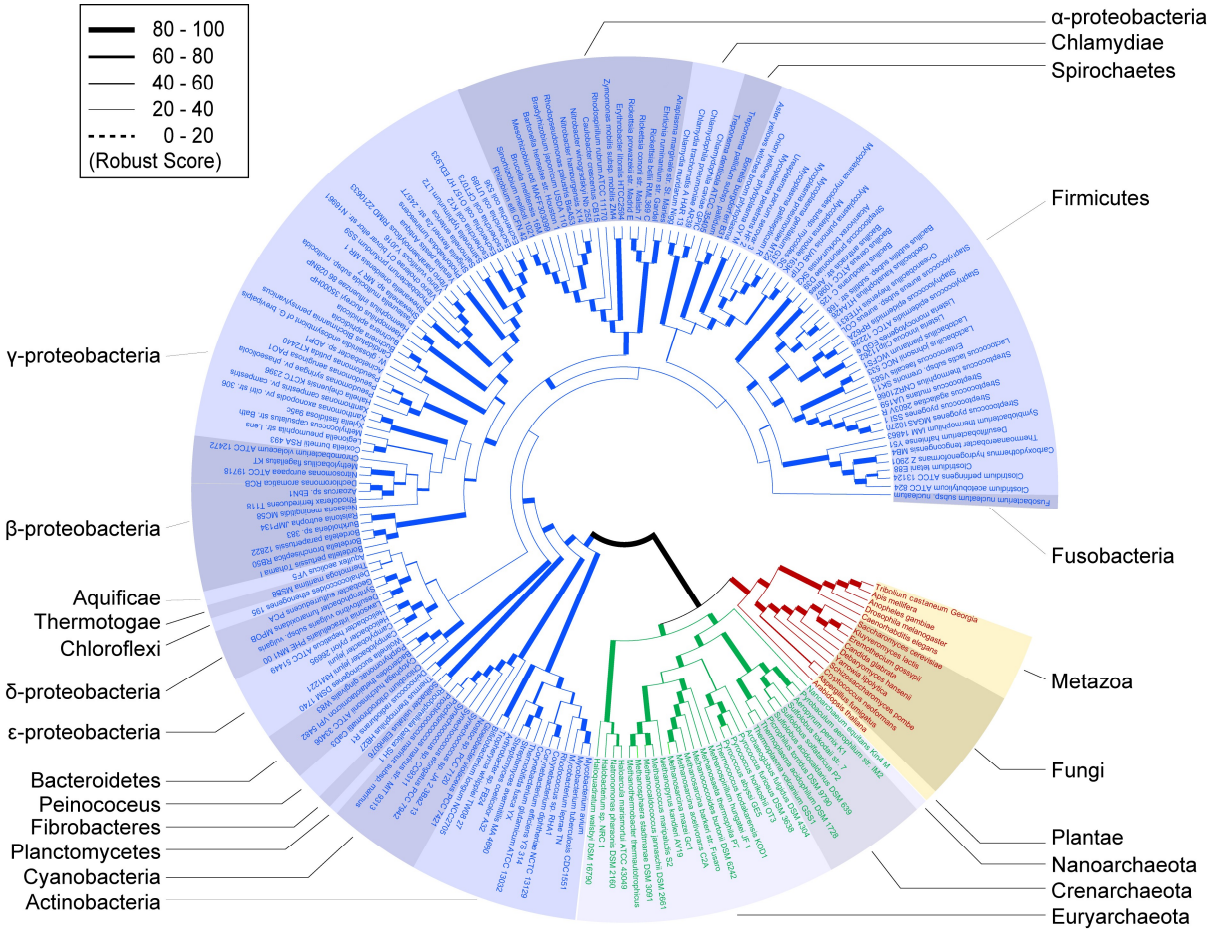


Figure 1.5. The "tree of life" based on genome context networks
This image is based on 195 genomes from representative species that had been fully sequenced by 2008 [62]. The small, red section shows the Eukaryota. The green and blue sections show the Archaea and Bacteria, respectively.

Even "The March of Progress" leading to modern humans has been updated. The formerly linear depictions of our origin have been replaced by more accurate branching depictions of hominin phylogeny, as shown below.

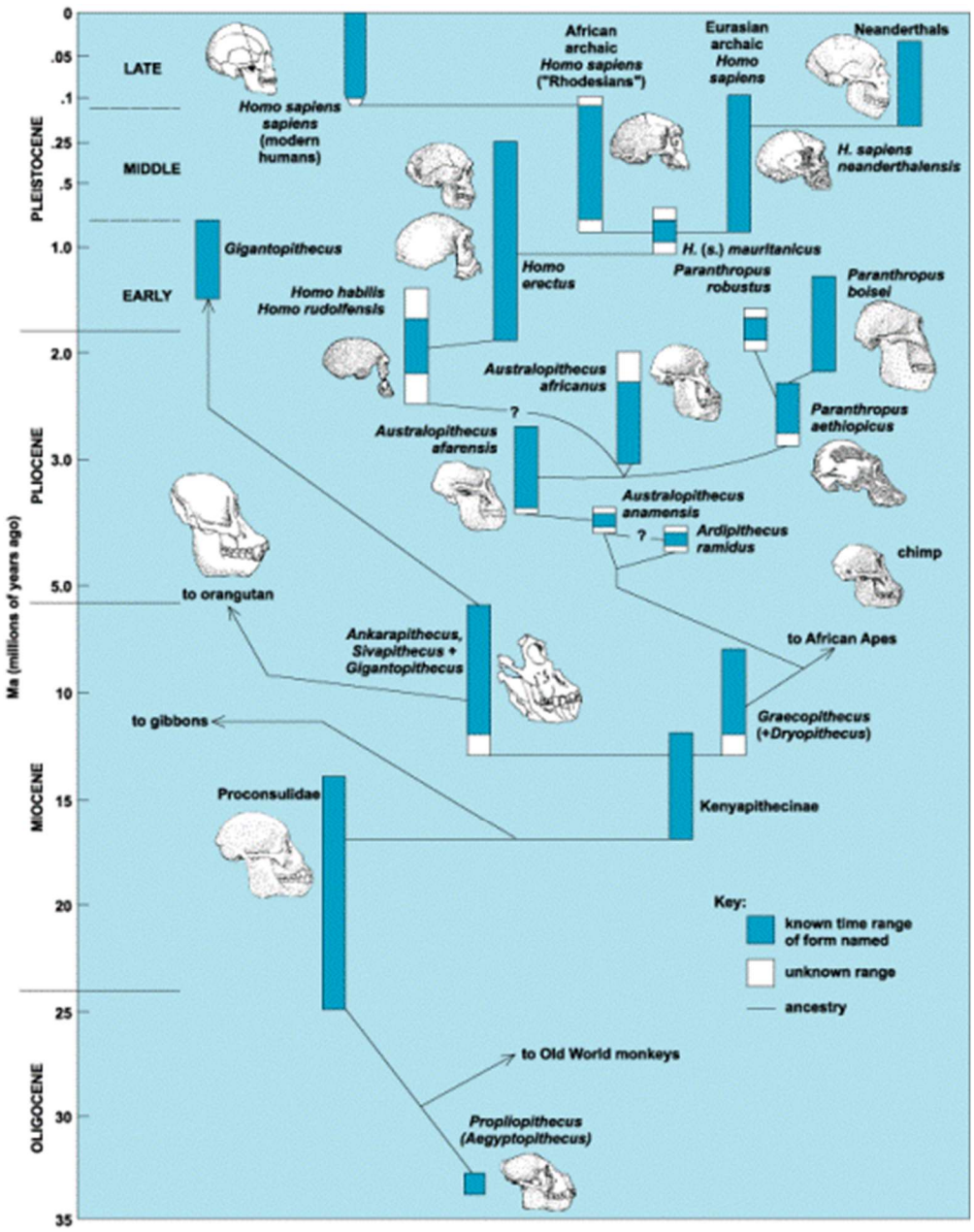


Figure 1.6. A modern depiction of hominin evolution
This branching design avoids a misleading linear progression to the emergence of modern humans [63].

Gould's idea of "replaying life's tape" offered a way, however fanciful, to assess—or at least imagine—the relative importance of the essential components of evolutionary change. Gould's thought experiment was a vehicle for him to express his ideas about the philosophical implications of historical contingency, including his perspective on human origins. Those

implications are still relevant today. To some, Gould's view is a "radical" challenge to anthropocentric, "fatalistic" reasoning about natural history. Gould said, "Replay the tape a million times from a Burgess beginning, and I doubt that anything like *Homo sapiens* would ever evolve again. It is, indeed, a wonderful life" [1].

The Experimental Era: Replaying Life's Tape in the Laboratory

Despite the seeming impracticality of Gould's original vision, an impressive empirical research program has taken up the mantle of his thought experiment [64]. This research includes experimental and comparative studies of the roles of contingency and convergence in evolution [57,65–77]. In particular, the rise of experimental evolution [78,79] was leveraged to transform Gould's idea into a method for conducting replay experiments using rapidly evolving microbes [64]. These experiments are not on the grand scale that Gould imagined in *Wonderful Life*. Nonetheless, replay experiments allow researchers to determine what happens when history is repeated in a laboratory. "Replaying life's tape" allows experimental evolutionary biologists to measure and compare the contributions of the essential components of evolution under various scenarios.

In a pioneering example, Travisano et al. designed and executed a two-phase "historical difference" experiment with bacteria that explicitly partitioned the relative contributions of adaptation, chance, and history in their response to an environmental change [80]. In the first phase of their experiment, 12 replicate populations of *E. coli* evolved independently for 2,000 generations in a glucose-limited medium [81–83]. The replicate populations in this phase evolved from the same ancestral state and under identical conditions. The design of the first phase thus generates simultaneous, rather than sequential, replays of evolution. The hallmark of Travisano et al.'s experimental design lies in the second phase [80]. A total of 36 clonal isolates, three replicates

of a single genotype from each population that evolved in the first phase, were used to found populations that then evolved for 1,000 generations in a new environment. In the new environment, maltose replaced the glucose used in the original environment. Following the second phase, Travisano et al. were able to measure the contributions of adaptation (change in the grand mean between the lines derived in the second phase and their proximate ancestors), chance (variation among lines derived from the same ancestor), and history (variation between groups of lines derived from different ancestors) to the evolution of fitness (a trait under strong directional selection) and cell size (a trait not subjected to direct selection) in the new, maltose-limited environment.

Similar experiments have now been conducted in a number of biological and digital systems [84–89]. These experiments have demonstrated that adaptation to new environments often overwhelms previously evolved historical differences [64]. However, most experiments have been conducted over relatively short periods of time and, in a review, Blount et al. noted that closely related lineages are "predisposed to evolve in the same way" [64]. In that vein, Travisano et al. had speculated that "over much longer periods, the footprint of history might eventually become too deep to be obscured even by intense selection [80]." Indeed, one recent study found that "contingency generated over long historical timescales steadily erased necessity" in the experimental evolution of ancestral proteins that diverged millions of years ago [90]. Measuring the effect of the "footprint of history" in digital organisms is the primary focus of the experiments and activity reported in the rest of my dissertation.

In the next chapter, we use the Avida platform to perform a two-phase historical difference experiment to estimate the effect of *deepening* the footprint of history on the evolution of digital organisms encountering new environments. In the first phase, we evolved ten replicate populations

from a common ancestor under identical conditions for tens of thousands of generations. The design of the first phase is similar to Lenski's Long-Term Evolution Experiment (LTEE) with *E. coli* [81,91,92]. We isolated the dominant genotype from each lineage at three different times. These times correspond to shallow, intermediate, and deep evolutionary histories in the ancestral environment. In the second phase, we evolved ten replicate populations from each of the 30 genotypes in two new environments. One environment was similar to the environment used during the first phase and the other was more novel. Following the second phase, we measured the contributions of adaptation, chance, and history to the evolution of fitness and genome length for each depth of history and in each new environment, using the statistical framework of Travisano et al. [80]. We thereby assessed how deepening the footprint of history in the initial environment shaped evolution during the second phase. This design also permitted us to determine whether the effects associated with deepening the footprint of history depended on the degree of similarity between the ancestral and novel environments.

In *Wonderful Life* Gould speculated, "perhaps genetic systems do 'age' in the sense of becoming 'less forgiving of major restructuring' … Perhaps modern organisms could not spawn a rapid array of fundamentally new designs, no matter what the ecological opportunity" [1]. Our analyses explore the effect of "genetic aging" on adaptation to new opportunities, and how that effect interacts with the degree of environmental similarity between old and new environments.

In the third chapter, we perform an extended analysis of the gain-of-function mutations that occurred during the second phase of the experiment reported in the preceding chapter. Following Ostrowski et al. [93], we determine the direct and indirect fitness effects of each gain-of-function mutation along the line of descent for the final dominant organism from each of the 600 populations. The purpose of this analysis is to determine how positive and negative pleiotropic

effects of mutations on fitness in the ancestral and novel environments contribute to the observed effects of the footprint of history.

The fourth and final chapter presents an educational activity. The exercise allows students to explore the concept of historical contingency by conducting the second phase of a historical-difference experiment and then analyzing the results with their classmates. We conducted the first phase of the two-phase experiment beforehand, both to reduce the time required for students to complete the exercise and to identify intermediate ancestors that would consistently provide clear outcomes. To that end. we evolved 15 replicate populations from the Avida-ED default ancestor in the same environment. From three of these populations, we identified genotypes that showed distinct evolutionary potentials when used as ancestors in the alternative environment that the students will use when performing the second phase. The students will be divided into three teams, and each team will use one of the three ancestors in their experiments. Each student will evolve 10 replicate populations from the ancestor assigned to their team. Next, the students on each team will work together to assess the performance of the replicates derived from their ancestor. Finally, the three teams will share data to compare the evolutionary performance of replicates derived from the three different ancestors. The different evolutionary tendencies of the three groups in the second phase reflect historical differences that arose while each ancestral lineage evolved during the first phase.

In conclusion, Gould's analysis of the "conventional view" and the associated "iconography of progress" shows how his own "radical" views challenge anthropocentric interpretations of natural history. While his forceful assertion of the pervasiveness of contingency in evolution is still debated, his thought experiment of "replaying life's tape" has inspired a growing body of research on historical effects in evolution, including my own dissertation.

REFERENCES

REFERENCES

1. Gould SJ. Wonderful Life. W. W. Norton; 1989.
2. Warren R. The Purpose Driven Life. Zondervan; 2012.
3. TED. Dan Dennett: Responding to Pastor Rick Warren. 2007.
4. Gould SJ. Evolution as fact and theory. Discover. 1981;2(5):34–37.
5. Valadés D. Rhetorica Christiana. 1579.
6. Blake K. On the Origins of "The March of Progress". Washington University; 2018.
7. Depew DJ. Contingency, Chance, and Randomness in Ancient, Medieval, and Modern Biology. In: Chance in Evolution. University of Chicago Press; 2016.
8. Dewey J. The influence of Darwin on philosophy and other essays in contemporary thought. SIU Press; 2007.
9. Gould SJ. Evolution Revolution. The Skeptics Society in California; 2000.
10. Buffon G-LL. comte de. "L'asne." Histoire naturelle, générale et particulière. 1753;377–403.
11. Buffon G-LL. The Epochs of Nature. Chicago, University of Chicago Press; 1778.
12. Paley W. Natural Theology. Gould and Lincoln; 1863.
13. de Lamarck JB. Philosophie Zoologique. Savy; 1873.
14. Redi F. Experiments on the Generation of Insects. Open Court; 1909.
15. Locy WA. Biology and Its Makers. Henry Holt; 1922.
16. Jordan R. Time and Contingency in St. Augustine. Rev Metaphys. 1955;8(3):394–417.
17. Aristotle A. On Generation and Corruption. Les Belles Lettres; 1966.
18. de Maillet B. Telliamed. Urbana: University of Illinois Press; 1968.
19. Sloan PR. The idea of racial degeneracy in Buffon's Histoire Naturelle. Studies in Eighteenth-Century Culture. 1974;3(1):293-321.
20. Richards RJ. The Meaning of Evolution. University of Chicago Press; 1992.
21. Aristotle. Physics Book II. 1994.

22. Plato. Timaeus. Hackett; 2000.

23. Aristotle. On the Parts of Animals. Oxford University Press; 2002.

24. Osler MJ. Divine Will and the Mechanical Philosophy. Cambridge University Press; 2004.

25. Lucretius TC, Munro HA. On the Nature of Things. Wilder; 2012.

26. Muller RA. Divine Will and Human Choice. Baker Academic; 2017.

27. Sloan P. Evolutionary Thought Before Darwin. In: The Stanford Encyclopedia of Philosophy. Stanford University; 2019.

28. Aquinas T. Summa Theologica. De Gruyter; 2021.

29. Ulett MA. Making the case for orthogenesis: The popularization of definitely directed evolution (1890–1926). Stud Hist Philos Sci Part C. 2014;45:124–132.

30. Darwin C. On the Various Contrivances by Which British and Foreign Orchids are Fertilized. Murray; 1862.

31. Peirce CS. Evolutionary love. The Monist. 1893;3(2):176–200.

32. Simpson GG. The Meaning of Evolution. Yale University Press; 1949.

33. Simpson GG. Evolutionary determinism and the fossil record. The Scientific Monthly. 1950; 71(4):262–267.

34. Darwin C. On the Origin of Species. 6th ed. Macmillan; 1962.

35. Hacking I. The Taming of Chance. Cambridge University Press; 1990.

36. Hodge J, Radick G. The place of Darwin's theories in the intellectual long run. The Cambridge Companion to Darwin. Cambridge University Press. 2009;246–273.

37. Darwin C. The Variation of Animals and Plants Under Domestication. Cambridge University Press; 2010.

38. Beatty J. Chance variation and evolutionary contingency: Darwin, Simpson, the Simpsons, and Gould. 2008.

39. Simpson GG. The role of the individual in evolution. J Wash Acad Sci. 1941;31(1):1–20.

40. Oster G, Alberch P. Evolution and bifuraction of developmental programs. Evolution. 1982;444–459.

41. Smith JM. Developmental constraints and evolution: A perspective from the Mountain Lake conference on development and evolution. Q Rev Biol. 1985;60(3):265–287.

42. Gould SJ, Woodruff DS. History as a cause of area effects: An illustration from Cerion on Great Inagua, Bahamas. Biological Journal of the Linnean Society. 1990;40(1):67–98.

43. Parker GA, Smith JM. Optimality theory in evolutionary biology. Nature. 1990;348(6296):27–33.

44. Williams G. Natural selection: Domains, levels, and challenges. Oxford University Press; 1992.

45. Mayr E. How to carry out the adaptationist program? Am Nat. 1983;121(3):324–334.

46. Gould SJ, Lewontin RC. The spandrels of San Marco and the Panglossian paradigm: A critique of the adaptationist programme. Proc R Soc Lond B Biol Sci. 1979;205(1161):581–598.

47. Zuckerkandl E, Pauling L. Molecules as documents of evolutionary history. J Theor Biol. 1965;8(2):357–366.

48. Kimura M. Evolutionary rate at the molecular level. Nature. 1968;217(5129):624–626.

49. King J, Jukes T. Non-darwinian evolution. Science. 1969;164(3881):788–798.

50. Lande R. Natural selection and random genetic drift in phenotypic evolution. Evolution. 1976;30(2):314–334.

51. Kimura M. The Neutral Theory of Molecular Evolution. Cambridge University Press; 1983.

52. Wright S. The roles of mutation, inbreeding, crossbreeding, and selection in evolution. 1932.

53. Wright S. Random Drift and the Shifting Balance Theory of Evolution. In: Mathematical Topics in Population Genetics. Springer; 1970.

54. Wright S. The shifting balance theory and macroevolution. Annu Rev Genet.1982;16(1):1–20.

55. Provine WB. Sewall Wright and Evolutionary Biology. University of Chicago Press; 1989.

56. Dennett DC. Darwin’s Dangerous Idea. Simon and Schuster; 1996.

57. Morris SC. Life’s Solution. Cambridge University Press; 2003.

58. Yochelson EL. Charles Doolittle Walcott, Paleontologist. Kent State University Press; 1998.

59. Marty C. Darwin on a godless creation: “It’s like confessing to a murder”. Scientific American. 2009.

60. Darwin Correspondence Project. Cambridge; 2021.

61. The Tree of Life: We are related to every living thing! Evogeneao; 2017.

62. Ding G, Yu Z, Zhao J, Wang Z, Li Y, Xing X, Wang C, Liu L, Li Y. Tree of life based on genome context networks. PloS one. 2008;3(10):e3357.

63. Says A. More from the blogosphere on the Ilert Fossils. Anthropology.net. 2007.

64. Blount ZD, Lenski RE, Losos JB. Contingency and determinism in evolution: Replaying life's tape. Science. 2018;362(6415):1–10.

65. Losos JB, Jackman TR, Larson A, de Queiroz K, Rodriguez-Schettino L. Contingency and determinism in replicated adaptive radiations of island lizards. Science. 1998;279(5359):2115–2118.

66. Teotónio H, Rose MR. Variation in the reversibility of evolution. Nature. 2000;408(6811):463–466.

67. Emerson SB. A macroevolutionary study of historical contingency in the fanged frogs of Southeast Asia. Biol J Linn Soc. 2001;73(1):139–151.

68. Vanhooydonck B, Irschick DJ. Is evolution predictable? Evolutionary relationships of divergence in ecology, performance and morphology in Old and New World lizard radiations. Topics in Functional and Ecological Vertebrate Morphology. 2002;191–204.

69. Joshi A, Castillo RB, Mueller LD. The contribution of ancestry, chance, and past and ongoing selection to adaptive evolution. J Genet. 2003;82(3):147–162.

70. Ortlund EA, Bridgham JT, Redinbo MR, Thornton JW. Crystal structure of an ancient protein: Evolution by conformational epistasis. Science. 2007;317(5844):1544–1548.

71. Keller SR, Taylor DR. History, chance and adaptation during biological invasion: Separating stochastic phenotypic evolution from response to selection. Ecol Lett. 2008;11(8):852–866.

72. Morris SC. Life's solution: what happens when we re-run the tape of life? Studies. 2008;97(386):205–217.

73. Dick MH, Lidgard S, Gordon DP, Mawatari SF. The origin of *Ascophoran bryozoans* was historically contingent but likely. Proc R Soc B. 2009;276(1670):3141–3148.

74. Morris SC. Evolution: like any other science it is predictable. Phil Trans R Soc B. 2010;365(1537):133–145.

75. Meyer JR, Dobias DT, Weitz JS, Barrick JE, Quick RT, Lenski RE. Repeatability and contingency in the evolution of a key innovation in phage lambda. Science. 2012;335(6067):428–432.

76. Harms MJ, Thornton JW. Historical contingency and its biophysical basis in glucocorticoid receptor evolution. Nature. 2014;512(7513):203–207.

77. Starr TN, Picton LK, Thornton JW. Alternative evolutionary histories in the sequence space of an ancient protein. Nature. 2017;549(7672):409–413.

78. Kawecki TJ, Lenski RE, Ebert D, Hollis B, Olivieri I, Whitlock MC. Experimental evolution. Trends in Ecology & Evolution. 2012;27(10):547–560.

79. Lenski RE. What is adaptation by natural selection? Perspectives of an experimental microbiologist. PLoS Genet. 2017;13(4):1–12.

80. Travisano M, Mongold JA, Bennett AF, Lenski RE. Experimental tests of the roles of adaptation, chance, and history in evolution. Science. 1995;267(5194):87–90.

81. Lenski RE, Rose MR, Simpson SC, Tadler SC. Long-term experimental evolution in *Escherichia coli*. I. Adaptation and divergence during 2,000 generations. Am Nat. 1991;138(6):1315–1341.

82. Vasi F, Travisano M, Lenski RE. Long-term experimental evolution in *Escherichia coli*. II. Changes in life-history traits during adaptation to a seasonal environment. Am Nat. 1994;144(3):432–456.

83. Travisano M, Vasi F, Lenski RE. Long-term experimental evolution in *Escherichia coli*. III. Variation among replicate populations in correlated responses to novel environments. Evolution. 1995;49(1):189–200.

84. Wagenaar DA, Adami C. Influence of chance, history, and adaptation on digital evolution. Artif Life. 2004;10(2):181–190.

85. Pérez-Zaballos FJ, Ortega-Mora LM, Álvarez-García G, Collantes-Fernández E, Navarro-Lozano V, García-Villada L, et al. Adaptation of *Neospora caninum* isolates to cell-culture changes: an argument in favor of its clonal population structure. J Parasitol. 2005;507–510.

86. Flores-Moya A, Costas E, López-Rodas V. Roles of adaptation, chance and history in the evolution of the dinoflagellate *Prorocentrum triestinum*. Naturwissenschaften. 2008;95(8):697–703.

87. Rouco M, López-Rodas V, Flores-Moya A, Costas E. Evolutionary changes in growth rate and toxin production in the cyanobacterium *Microcystis aeruginosa* under a scenario of eutrophication and temperature increase. Microb Ecol. 2011;62(2):265–273.

88. Flores-Moya A, Rouco M, García-Sánchez MJ, García-Balboa C, González R, Costas E, et al. Effects of adaptation, chance, and history on the evolution of the toxic dinoflagellate *Alexandrium minutum* under selection of increased temperature and acidification. Ecol Evol. 2012;2(6):1251–1259.

89. Rebolleda-Gómez M, Travisano M. Adaptation, chance, and history in experimental evolution reversals to unicellularity. Evolution. 2019;73(1):73–83.

90. Xie VC, Pu J, Metzger BP, Thornton JW, Dickinson BC. Contingency and chance erase necessity in the experimental evolution of ancestral proteins. eLife. 2021;10:e67336.

91. Lenski RE, Travisano M. Dynamics of adaptation and diversification: A 10,000-generation experiment with bacterial populations. Proc Natl Acad Sci. 1994;91(15):6808–6814.

92. Tenaillon O, Barrick JE, Ribeck N, Deatherage DE, Blanchard JL, Dasgupta A, et al. Tempo and mode of genome evolution in a 50,000-generation experiment. Nature. 2016;536(7615):165–170.

93. Ostrowski EA, Ofria C, Lenski RE. Ecological specialization and adaptive decay in digital organisms. Am Nat. 2007;169(1):E1–20.

CHAPTER 2: HOW THE FOOTPRINT OF HISTORY SHAPES THE EVOLUTION OF DIGITAL ORGANISMS

Authors: Jason Nyerere Bundy, Charles Ofria, and Richard E. Lenski

Abstract

Gould's thought experiment of "replaying life's tape" provides a conceptual framework for experiments that quantify the contributions of adaptation, chance, and history to evolutionary outcomes. For example, we can empirically measure how varying the depth of history in one environment influences subsequent evolution in a new environment. Can this "footprint of history"—the genomic legacy of prior adaptation—grow too deep to overcome? Can it constrain adaptation, even with intense selection in the new environment? We investigated these questions using digital organisms. Specifically, we evolved ten populations from one ancestor under identical conditions. We then replayed evolution from three time points in each population's history (corresponding to shallow, intermediate, and deep history) in two new environments (one similar and one dissimilar to the prior environment). We measured the contributions of adaptation, chance, and history to the among-lineage variation in fitness and genome length in both new environments. In both environments, variation in genome length depended largely on history and chance, not adaptation, indicating weak selection. By contrast, adaptation, chance, and history all contributed to variation in fitness. Crucially, whether the depth of history affected adaptation depended on the environment. When the ancestral and new environments overlapped, history was as important as adaptation to the fitness achieved in the new environment for the populations with the deepest history. However, when the ancestral and novel environments favored different traits, adaptation overwhelmed even deep history. This experimental design for assessing the influence of the depth of history is promising for both biological and digital systems.

Introduction

In his seminal book *Wonderful Life* [1], Stephen Jay Gould argued for the importance of historical contingencies in the evolution of life on Earth. He proposed a sublime thought experiment of "replaying life's tape" to ask how similar the living world would be to what we see today if evolution was repeated from, say, the time of the Cambrian explosion. Gould argued that living organisms would likely be very different. However, he concluded, "The bad news is that we can't possibly perform the experiment."

But let us presume for a moment that we *could* perform the experiment. Why not up the ante? Why should we rewind the tape to a single point in time, such as the Cambrian explosion some 540 million years ago (mya)? While the *VHS Time Machine* is up and running, why not also dial in some replays from the Paleozoic/Mesozoic boundary (~250 mya) and the late-Cretaceous mass extinction (~65 mya)? By having replays from the dawn of different eras, we could study the influence of varying the time scale over which the signatures of historical contingency become manifest. We could then measure the "footprint of history." Our own thought experiment thus poses the additional question: "How different would the outcomes be if we rewound the tape of life to different points in history?"

Of course, there was more at stake in Gould's thought-experiment than biological fantasy. Gould's idea of "replaying life's tape" offered a way, however fanciful, to assess the relative importance of the essential components of evolutionary change: selection (adaptation), stochastic variation (chance), and phyletic inheritance (history). So, despite its seeming impracticality, an impressive empirical research program has taken up the mantle of Gould's *gedankenexperiment* [2]. This program includes experimental and comparative studies of the role of contingency and convergence in evolution [3–16]. In particular, the rise of experimental evolution has transformed

the idea of "replaying life's tape" from metaphor to methodology, allowing replay experiments to be conducted in the laboratory using rapidly evolving microbes [2]. While these experiments are not on the grand scale that Gould imagined in *Wonderful Life*, they nonetheless provide a way to estimate the relative contributions of these essential evolutionary components.

Of particular note, Travisano et al. [3] performed a two-phase experiment with bacteria that explicitly partitioned the relative contributions of adaptation, chance, and history to evolutionary change. In the first phase of their experiment, 12 replicate populations of *E. coli* evolved independently for 2,000 generations in a glucose-limited medium [17]. The populations in this phase evolved from the same initial state under identical conditions, thus serving as simultaneous, rather than sequential, replays. The hallmark of Travisano et al.'s experimental design lies in the second phase [3]. A total of 36 clonal isolates (three derived from each replicate in the first phase) evolved for 1,000 generations in a new environment, in which maltose replaced the glucose used in the original environment. Following the second phase, the authors were able to partition the contributions of adaptation (change in the grand mean between the derived Phase II lines and their proximate ancestors), chance (variation among Phase II lines derived from the same ancestor), and history (variation among Phase II lines with different ancestors) to the evolution of both fitness (a trait under strong directional selection) and cell size (a trait not subjected to direct selection) in the new, maltose-limited environment (Figure 2.1).

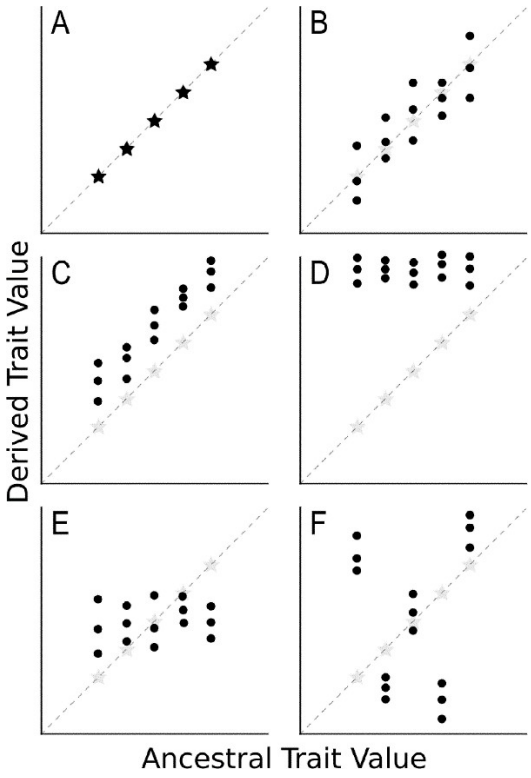


Figure 2.1. Schematic illustration of the contributions of adaptation, chance, and history
We imagine 5 ancestors arranged along the x-axis according to the mean value of some trait. Each ancestor founds 3 experimental populations that evolve in a new environment, and the mean trait values of the derived populations are shown along the y-axis. **(A)** Initial state: The contribution of history reflects the extent of spread along the x-axis. **(B)** Chance with history sustained: Random factors cause the sets of 3 derived populations that share the same ancestor to diverge from one another, but without any systematic directional tendency. **(C)** Adaptation with history sustained and minimal effect of chance: The trait values of all 15 derived populations have increased (or decreased) systematically relative to their ancestors, and to a similar degree regardless of the ancestral state. **(D)** Adaptation with history erased: The trait values of all 15 derived populations have increased relative to their ancestors, and to a greater extent in those populations founded by ancestors with lower values, such that the effect of history has been eliminated. **(E)** History diminished without directional trend: This pattern might reflect random factors obscuring the effect of history or, alternatively, stabilizing selection on an intermediate trait value. **(F)** History amplified idiosyncratically: Sets of derived populations vary consistently depending on their ancestors, but without any directional trend in trait values or systematic correlation of derived and ancestral values. Modified from Travisano et al. [3].

Experiments of a similar design have now been conducted using dinoflagellates [18,19], an apicomplexan parasite [20], a cyanobacterium [21], the brewer's yeast *Saccharomyces cerevisiae* [22], and the bacterial pathogen Acinetobacter baumannii [23]. These studies have

shown that adaptation typically dominates the evolution of fitness and related traits under intense directional selection, whereas chance and history play greater roles for traits subject to little or no selection. As fitness and other strongly selected traits evolve, the effects of the Phase II lineages' divergent histories—driven by the effects of chance, adaptation, or both in the Phase I environment—were often largely erased [2,3,24–29]. Digital organisms—computer programs that replicate, mutate, compete, and evolve in virtual environments—have also been used to analyze contingency in evolving systems [30–35]. The results of these digital experiments have been consistent with those performed using microorganisms, with adaptation dominating the evolution of traits closely tied to fitness, while chance and history play more prominent roles in the evolution of other traits.

Of particular relevance for our work, Travisano et al. [3] concluded their paper with an intriguing speculation about the possible effect of extending the duration of history in the ancestral environment. Despite having shown that 1,000 generations in maltose (Phase II) largely erased the divergence that took place during 2,000 generations in glucose (Phase I), they suggested: "Over much longer periods, the footprint of history might eventually become too deep to be obscured even by intense selection." Comparative evidence supports the idea that convergent evolution is less likely when starting from more deeply divergent states [2]. To our knowledge, however, no experiment has systematically quantified the effects of varying the duration of evolution in an ancestral environment—in essence, varying the depth of the footprint of history. Our study addresses this limitation by means of appropriately designed experiments using digital organisms.

Experimental Design

We performed a set of two-phase replay experiments to test the effects of varying the depth of history in an ancestral environment on subsequent evolution in a new environment. We ran our

experiments using the Avida platform for digital evolution [32,36–38]. In Phase I, we started 10 populations from a single proto-ancestral clone that could self-replicate, but which could not perform any of the logic functions that may increase growth rates in their virtual world. These 10 populations evolved independently, but all in the same ancestral ("old") environment, for 500,000 updates (the basic time unit in Avida), which allowed tens of thousands of generations (mean = 66,074 generations). In this ancestral environment, digital organisms received resources that increased their growth rates if they performed any of 38 distinct two- or three-input logic functions. We isolated the most abundant genotype from each population at three time points, representing three depths of history: 20,000 updates ("shallow"), 100,000 updates ("intermediate"), and 500,000 updates ("deep"). These 30 clones (10 populations x 3 depths of history) were used to seed the Phase II populations.

In Phase II, each of the 30 proximate ancestors identified in Phase I founded 10 new populations in each of 2 new environments, for a grand total of 600 populations (30 clones x 2 environments x 10 replicates). We evolved each of these populations for 100,000 updates, which was equivalent to the intermediate depth in Phase I. One Phase II environment, which we call "Overlapping", provides the digital organisms with additional resources for performing 76 two- and three-input logic functions, including the 38 functions rewarded in the old environment as well as 38 additional functions. The other Phase II environment, called "Orthogonal", rewards those 38 additional, novel functions but does not reward the functions that were available in the ancestral environment. We then used the analytical framework of Travisano et al. [3] to calculate the contributions of adaptation, chance, and history to the evolution of average fitness and genome length in each new environment. We are particularly interested in examining the effects of

deepening the footprint of history on the relative contributions of those factors, and in testing whether the effects differ between the two new environments (Figure 2.2).

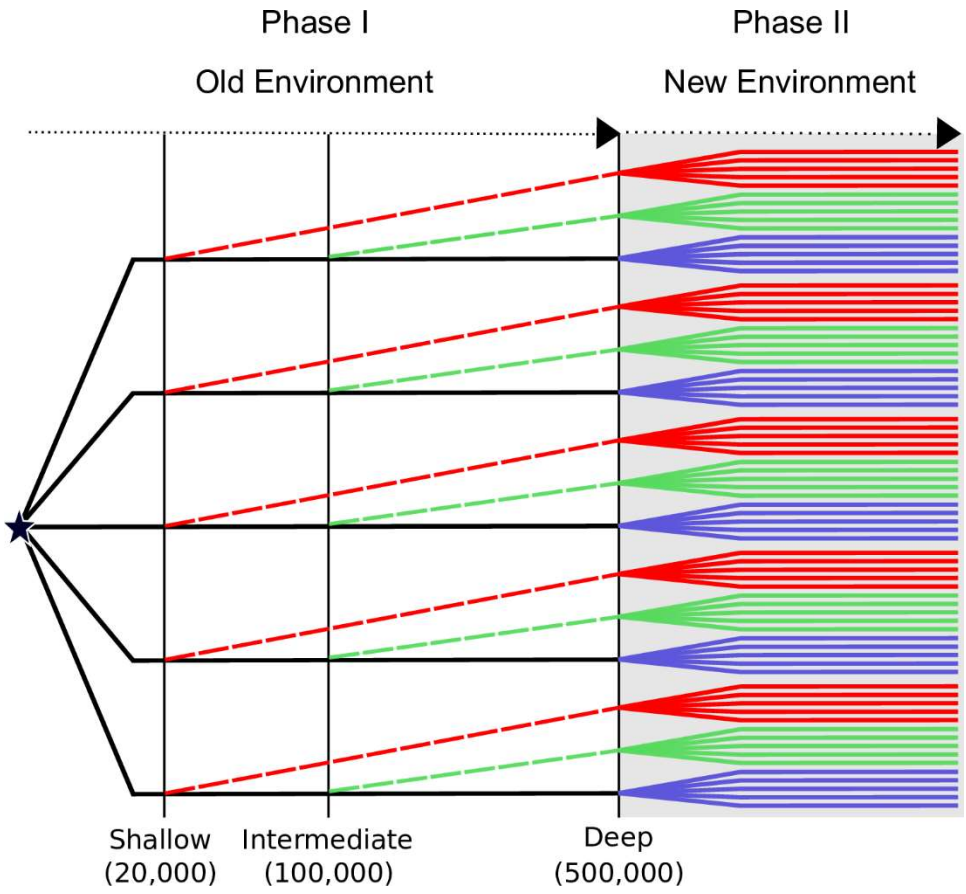


Figure 2.2. Schematic illustration of the experimental design
In Phase I, ten populations (five depicted) evolved from a single proto-ancestor (star) for 500,000 updates. We sampled the most abundant genotype from each population after 20,000, 100,000, and 500,000 updates, representing shallow, intermediate, and deep history, respectively. In Phase II, each of these 30 proximate ancestors founded ten replicate populations (five depicted) in each of two new environments (one depicted), where they then evolved for 100,000 updates.

In sum, we track 300 evolving populations in each of two environments. Each set of 300 populations has three cohorts (each 100 populations) defined by the depth of their history, with ten founders in each cohort and ten replicate populations per founder. Within each cohort, significant differences in the overall average trait value between replicates (at the end of Phase II) and their direct ancestors (at the outset of Phase II) reflect *adaptation*. For each group of ten replicate

populations, evolution proceeds from the same starting point under identical conditions, with only *chance* to explain variation among the replicates. Each lineage's unique *history* during Phase I accounts for any significant differences among the groups with different founders. Lastly, the *depth of the historical footprint*—reflecting the duration of evolution in the ancestral environment—is responsible for any significant differences among the three cohorts. By employing the analytical methods of Travisano et al. [3], we can estimate the relative contributions of adaptation (systematic changes in mean trait values), chance (variation among replicates), and history (variation among groups with different founders within a cohort). We then compare the contributions among the shallow, intermediate, and deep cohorts to assess the effect of varying the depth of history. We also compare results from the Overlapping and Orthogonal environments to explore whether the impact of that historical footprint depends on the degree of novelty imposed by the new environment.

Results

Evolution in Phase I

Starting from the same proto-ancestral genotype, all 10 populations underwent extensive evolution during Phase I. Mean fitness rose rapidly at first, and more slowly as time progressed (Figure 2.3). The populations all improved relative to their common ancestor, but they reached very different fitness levels (note the $\log_{10}$ scale), with the among-population variation increasing over time. The vertical lines show the time points (20,000, 100,000, and 500,000 updates) at which we sampled individual genotypes to start the Phase II populations. We also measured average genome length in each population. Although some variation in genome length emerged over time, there was no consistent directional change: 7 populations finished with slightly longer genomes than the ancestor, 3 ended with somewhat shorter genomes, and the direction of change in an individual population often shifted over time. Moreover, the changes in genome length and the variation

among the evolving populations were very small when compared to the fitness trajectories (keeping in mind the different scales for fitness and genome length). Taken together, these observations indicate that genome length was, if not strictly neutral, at least subject to much weaker selection than fitness.

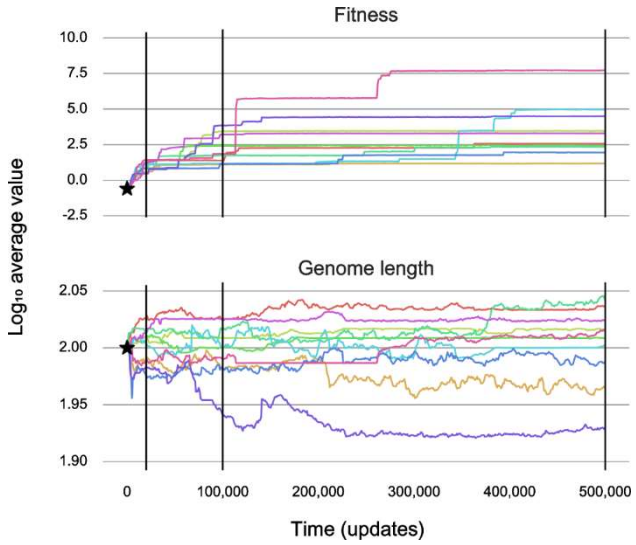


Figure 2.3. Population averages for fitness and genome length during Phase I
Average fitness (top) and genome length (bottom) measured every 1,000 updates for 10 replicate populations derived from a single proto-ancestor (shown by star), as they evolved in the same Phase I environment for 500,000 updates. We log-transformed both fitness and genome length. Vertical lines at 20,000, 100,000, and 500,000 updates show the times at which we chose the most abundant genotype in each population to start the Phase II replicates.

Phase II: Evolution in the Overlapping environment after a shallow Phase I history

Trajectories for fitness and genome length Figure 2.4 (top) shows the mean-fitness trajectories for 100 populations in the Overlapping environment, following a shallow history of only 20,000 updates during Phase I. In this environment, the organisms continued to receive rewards for performing the 38 logic functions rewarded in Phase I. This environment also rewarded the performance of 38 additional functions. We plotted 10 sets of 10 trajectories each, colored according to their Phase I progenitors. Figure 2.4 (bottom) shows the trajectories for average genome length in the same 100 populations. As in Phase I, fitness increased in all the Phase II

populations, whereas genome length hardly changed in most populations and showed little evidence of directionality overall.

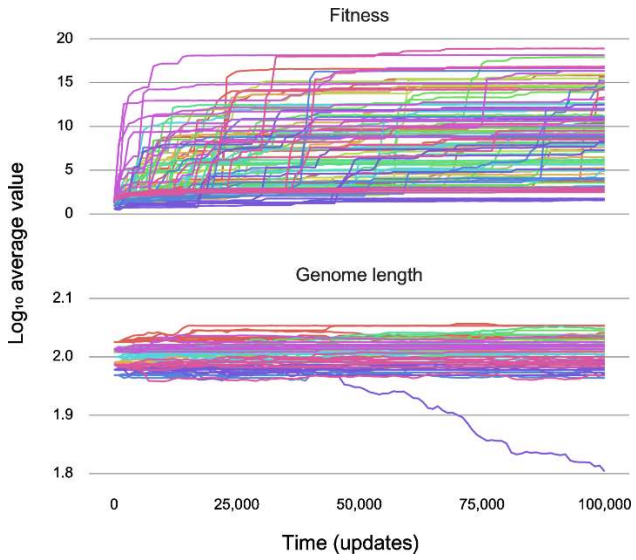


Figure 2.4. Population averages for fitness and genome length during Phase II in the Overlapping environment for populations with a shallow history
Average fitness (top) and genome length (bottom) measured every 1,000 updates for 10 replicate populations derived from a single founder from each Phase I population, as they evolved for 100,000 updates in the new Phase II environment. These populations have a shallow history, as their founders evolved for only 20,000 updates in the ancestral environment. The new environment rewarded performance of the same 38 logic functions as the old environment along with a suite of 38 additional logic functions.

Contributions of adaptation, chance, and shallow history We can use the data from the endpoints of the evolutionary trajectories to calculate the contributions of adaptation, chance, and history to each phenotype, using the same statistical approach as Travisano et al. [3]. Recall that adaptation represents the overall, directional shift in a trait across populations during Phase II; chance reflects the variation that arose during Phase II among replicates derived from the same Phase I progenitor; and history reflects the variation associated with the different Phase I progenitors. Figure 2.5 shows these contributions for fitness and genome length after 100,000 updates during Phase II, starting from the 10 progenitors that had evolved for only 20,000 updates

in Phase I, thus reflecting a shallow footprint of history. Both panels also show the ancestral variation in those traits among the 10 different progenitors.

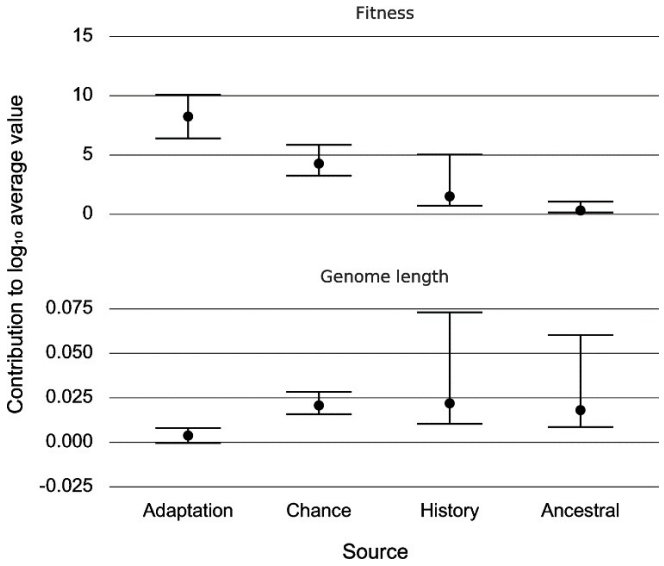


Figure 2.5. End-of-experiment effects of adaptation, chance, and history in the Overlapping environment with shallow history
Contributions of adaptation, chance, and history to the evolution of fitness (top) and genome length (bottom) based on end-of-experiment average values for populations with a shallow history in Phase I that evolved in the Overlapping environment during Phase II. Error bars show 95% confidence limits for the associated contributions, along with the ancestral variation at the start of Phase II.

In the case of fitness (Figure 2.5 top), we see a large contribution of adaptation, which reflects the systematic improvement in all 100 Phase II populations. We also see significant, but smaller, contributions that reflect the effects of chance and history. All three contributions, individually and collectively, are greater than the ancestral variation that existed at the start of Phase II. This ancestral variation is the initial Phase II historical variation, which arose during Phase I. For genome length (Figure 2.5 bottom), by contrast, we see no significant contribution of adaptation (the confidence interval includes 0). However, both chance and history contributed significantly to the final genome lengths, and the variation due to history is similar in magnitude to the ancestral variation in that trait.

We performed these analyses using the values at the end of Phase II, but we also measured the trajectories for all three contributors over time, as shown in Figure 2.6. We see the rising impact of adaptation on fitness over time, as well as the more slowly rising and meandering contributions of chance and history (Figure 2.6 top). In the case of genome length, we see the growing impact of chance events, a near-constant effect of history, and a much smaller contribution of adaptation (Figure 2.6 bottom).

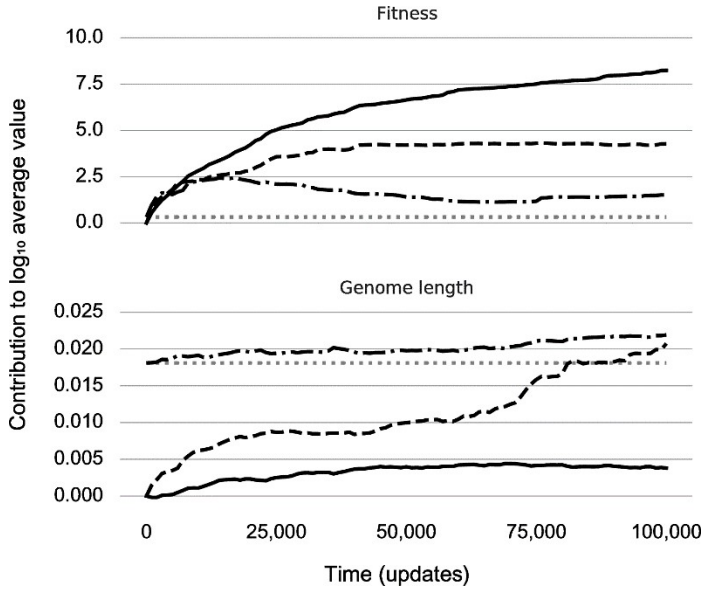


Figure 2.6. Trajectories of adaptation, chance, and history in the Overlapping environment with shallow history
Contributions of adaptation (solid line), chance (dashed line), and history (dash-dotted line) to the evolution of fitness (top) and genome length (bottom) measured every 1,000 updates during Phase II. The plot includes the ancestral variance (dotted line) for comparison.

Phase II: Evolution in the Orthogonal environment after a shallow Phase I history

We now turn to the results of evolution in the Orthogonal environment, using the same 10 progenitors that evolved for only 20,000 generations in Phase I. From the perspective of the digital organisms, the Orthogonal environment is more novel than the Overlapping environment. The Overlapping environment continued to reward organisms for performing the same 38 logic functions that the ancestral environment rewarded during Phase I. It also rewarded 38 additional

functions exclusive to Phase II. By contrast, the Orthogonal environment rewarded *only* these new functions, and not those rewarded in the ancestral environment. In both new environments, some genetic instructions involved with performing the previously evolved functions might serve as building blocks that could be repurposed to perform the newly rewarded functions. Having led the reader carefully through the data and inferences for the Overlapping environment, we relegate most data and analyses for the Orthogonal environment to the Supporting Information.

Trajectories for fitness and genome length The fitness trajectories in the Orthogonal environment (Figure 2.10 top) are similar in form to those seen in the Overlapping environment (Figure 2.4 top). However, the trajectories in the Orthogonal environment started from lower values, and typically reach lower final values, than in the Overlapping environment. The differences between the fitness and genome length trajectories according to environment reflect the continuity of selection. Adaptations acquired during Phase I carried over to the Overlapping environment. However, fitness collapsed in the transition to the Orthogonal environment because the ancestors from Phase I generally lacked the ability to perform functions rewarded in this more distinct, new environment. The trajectories for average genome length in the Orthogonal environment (Figure 2.10 bottom) were also similar to the corresponding trajectories from the Overlapping environment (Figure 2.4 bottom). In both cases, most trajectories showed little net directional change, indicative of a trait that is nearly neutral.

Contributions of adaptation, chance, and shallow history As before, we can use the final trait values to calculate the contributions of adaptation, chance, and history to mean fitness and genome length (Figure 2.11), as well as the temporal trajectories for these contributions (Figure 2.12). For genome length, the contributions of adaptation, chance, and history were nearly indistinguishable in the two environments. In both environments, the lingering effects of history hardly changed

over time, while the effects of chance crept upward, slightly eclipsing the effect of history in the Orthogonal environment. Also, adaptation had no significant effect on genome length in either environment. These trajectories indicate a trait that experiences little or no directional selection. The contributions of adaptation, chance, and history to fitness were also similar in the two environments, and their rank orders are the same. However, all the values are substantially larger in the Overlapping environment. Again, this difference in scaling reflects the fact that additional logic functions continued to be rewarded in the Overlapping environment, leading to higher trajectories there (Figure 2.4 top) than in the Orthogonal environment (Figure 2.10 top).

Summarizing to this point, we see a clear difference between the contributions of adaptation, chance, and history to the two traits that we measured. For fitness itself, adaptation had the largest effect, followed by chance producing substantial divergence even among the replicate populations founded from the same proximate ancestor. The lingering effect of history, reflecting the different proximate ancestors, was also significant but smaller than the other two effects. For genome length, by contrast, the effects of chance and history were similar in magnitude and both significant, whereas the effect of adaptation, if any, was small and not significant. Also, it would seem thus far that the environment hardly matters for these patterns. We now turn our attention to whether and how these results depend on the depth of the footprint of history.

Effects of deepening the footprint of Phase I history on evolution during Phase II

The results above are for Phase II populations founded by progenitors that had diverged from the common proto-ancestor for only 20,000 updates. In other words, the footprint of history during Phase I was shallow relative to the duration of Phase II, which was 100,000 updates. Here we examine the effects of history by also using genotypes sampled after 100,000 and 500,000 updates of Phase I to start populations, which then evolved for 100,000 updates in Phase II. In comparison

to the shallow footprint of history generated during 20,000 updates, we refer to the footprints that emerged after 100,000 and 500,000 updates as intermediate and deep, respectively.

We provide the underlying data in the Supporting Information, as follows. Figure 2.13 shows the trajectories for mean fitness in the Overlapping and Orthogonal environments, in each case using 10 founders sampled from Phase I after 100,000 updates (Figure 2.3). Figure 2.14 shows the corresponding trajectories for genome length in the two environments. The data in Figures 2.13 and 2.14 thus reflect the intermediate footprint of history. Figure 2.15 shows fitness trajectories in the Overlapping and Orthogonal environments, using 10 founders sampled from Phase I after 500,000 updates; and Figure 2.16 shows trajectories for genome length in these environments. The data in Figures 2.15 and 2.16 reflect the deep footprint of history.

In the two sections below, we examine the contributions of adaptation, chance, and history to genome length and fitness, respectively, comparing the contributions as a function of the depth of history during Phase I and the selective environment in Phase II. We invert the order of presentation and begin with genome length because the patterns and interpretation are simpler than they are for fitness.

Effect of deepening the footprint of history on genome length Figure 2.7 (top) shows the contributions of adaptation, chance, and history to final genome length as a function of the depth of history during Phase I for the populations that evolved in the Overlapping environment. On the right, one sees the effect of history on the ancestors used to found the Phase II populations; this effect became progressively larger as the depth of history increased from 20,000 to 100,000 to 500,000 updates. To the left of the ancestral estimates, one sees that these historical effects were maintained in the evolving lineages during Phase II. Chance shows the opposite trend, with its contribution declining as the depth of history increased. Adaptation—as indicated by consistent

directional changes in trait values—contributed little to genome length in comparison to either history or chance. However, the insignificant contribution of adaptation to genome length in the Overlapping environment when the depth of history was shallow (20,000 updates) became significant when history was deeper (100,000 and 500,000 updates), although the effect remains small. Figure 2.17 (top) shows the temporal trends in the contributions of adaptation, chance, and history to genome length during Phase II for all three depths of history.

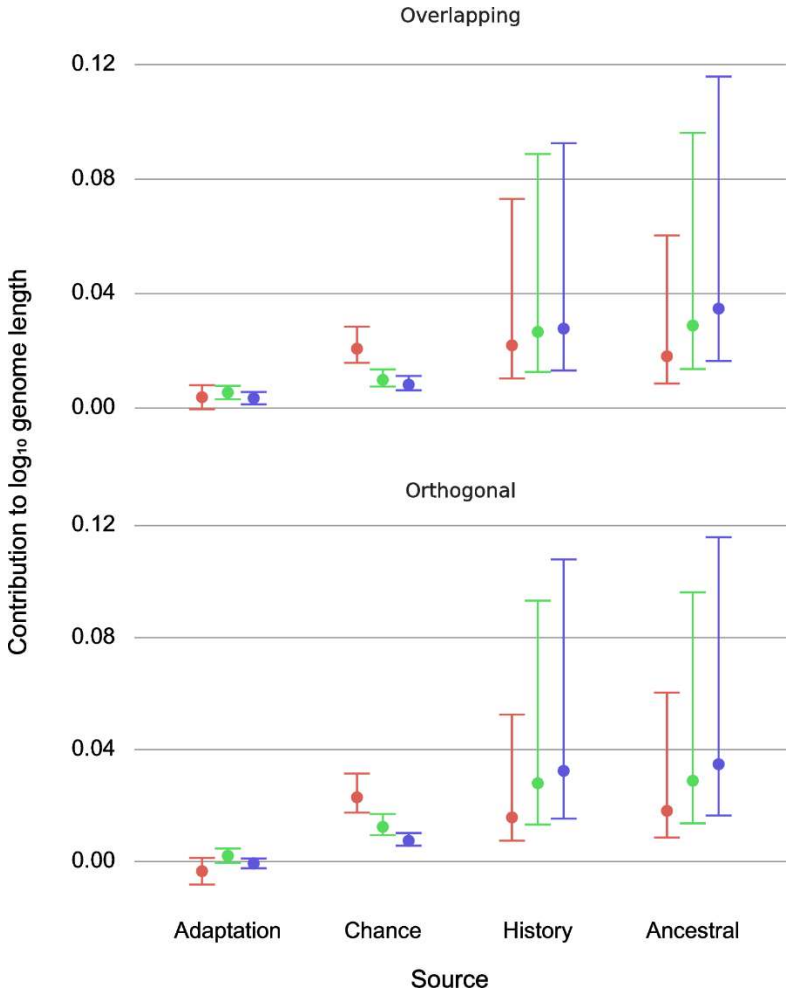


Figure 2.7. Impact of the footprint of history on the contributions of adaptation, chance, and history to genome length is similar in the two Phase II environments
Contributions of adaptation, chance, and history to genome length for populations with shallow (red), intermediate (green), and deep (blue) Phase I histories. Contributions are based on end-of-experiment averages for populations that evolved in the Overlapping (top) and Orthogonal (bottom) Phase II environments. Error bars show 95% confidence limits. The plot shows ancestral variation at the start of Phase II for comparison.

The end-of-experiment estimates and their trajectories are similar when we look at the effects of the depth of history on genome length in the Orthogonal environment (Figures 2.7 bottom and 2.17 bottom). The ancestral divergence increases with the depth of history, as do the persistent contributions of history, while the effect of chance declines as the footprint of history gets deeper. The only qualitative difference in the Orthogonal environment, when compared to the

Overlapping regime, is that there was no significant adaptation associated with the deeper histories. Given the lack of continuity in selection for specific logic functions, even the weak selection on genome length that might have occurred in the Overlapping environment was evidently inconsequential in the Orthogonal environment.

Effect of deepening the footprint of history on fitness The results for fitness are more complex, with populations in the two environments showing different effects of the depth of history. Figure 2.8 shows the contributions of adaptation, chance, and history to final mean fitness for the three depths of history in the Overlapping (top) and Orthogonal (bottom) treatments. In the Overlapping environment, adaptation swamped the effect of history when the footprint of history was shallow. As the depth of history increased from 20,000 to 100,000 and 500,000 updates, the contribution of adaptation declined, while the contribution of history to fitness increased. In fact, for populations with the deepest ancestry, history's contribution was as great as or slightly greater than that of adaptation, whereas for populations with the shallowest history the contribution of adaptation was several times that of history.

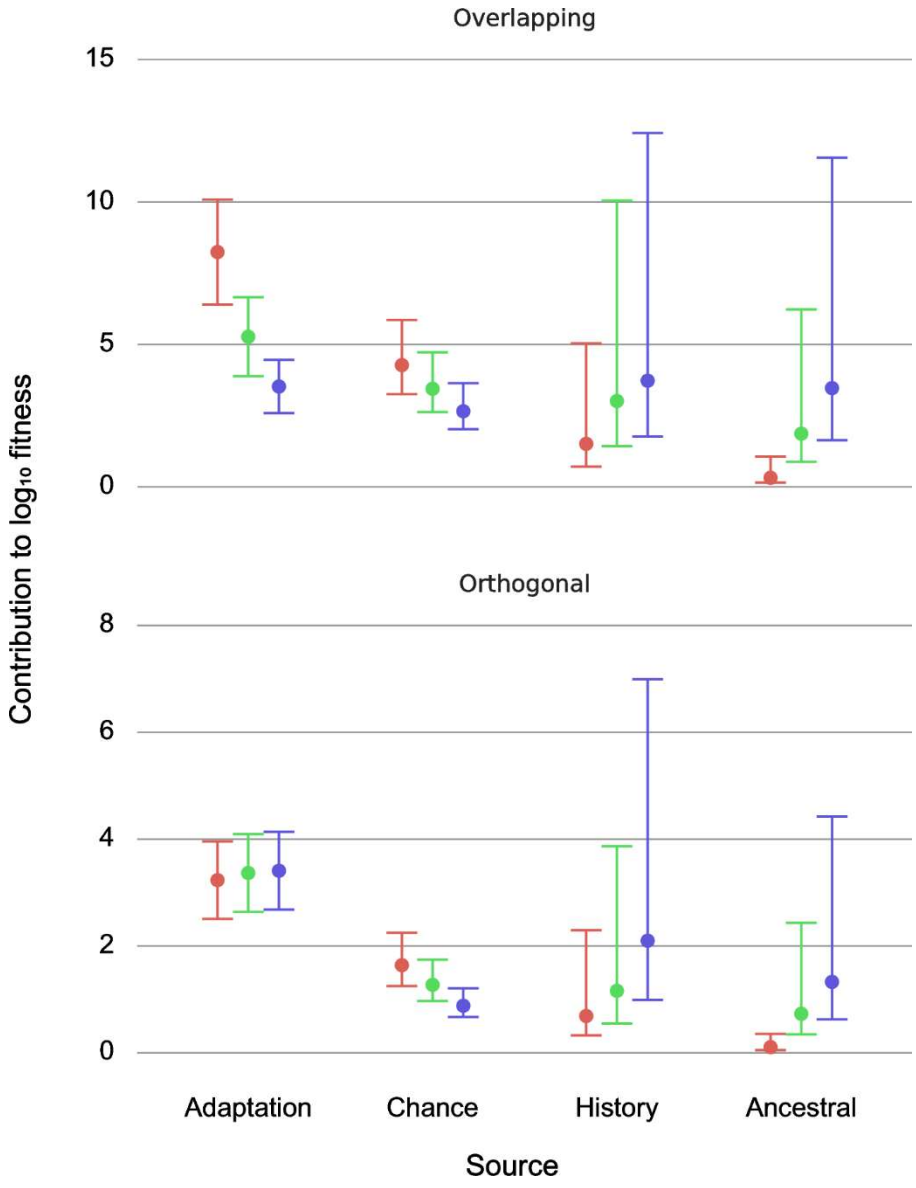


Figure 2.8. Impact of the footprint of history on the contributions of adaptation, chance, and history to fitness differs between the two Phase II environments
Contributions of adaptation, chance, and history to mean fitness for populations with shallow (red), intermediate (green), and deep (blue) Phase I histories. Contributions are based on end-of-experiment values for populations that evolved in the Overlapping (top) and Orthogonal (bottom) Phase II environments. Error bars show 95% confidence limits. The plot shows ancestral variation at the start of Phase II for comparison.

By contrast, in the Orthogonal environment, where the logic functions rewarded in Phase I were no longer rewarded during Phase II, the contribution of adaptation was essentially independent of the depth of history. The contribution of history to fitness nonetheless increased

with its depth, as it did in the Overlapping environment. However, even when the Phase I history was its deepest at 500,000 updates, and thus 5-fold longer than Phase II, the contribution of adaptation to fitness in the new environment remained greater than that of history.

Figure 2.9 shows the temporal trajectories of the estimated contributions of adaptation, chance, and history during Phase II in the Overlapping and Orthogonal environments. These trajectories further illustrate the impact of the different environments on the relative impact of adaptation and history. Notice that in the Orthogonal environment, the contribution of adaptation surpassed that of even the deepest history within a few thousand updates, whereas it had not done so even after 100,000 updates in the Overlapping environment.

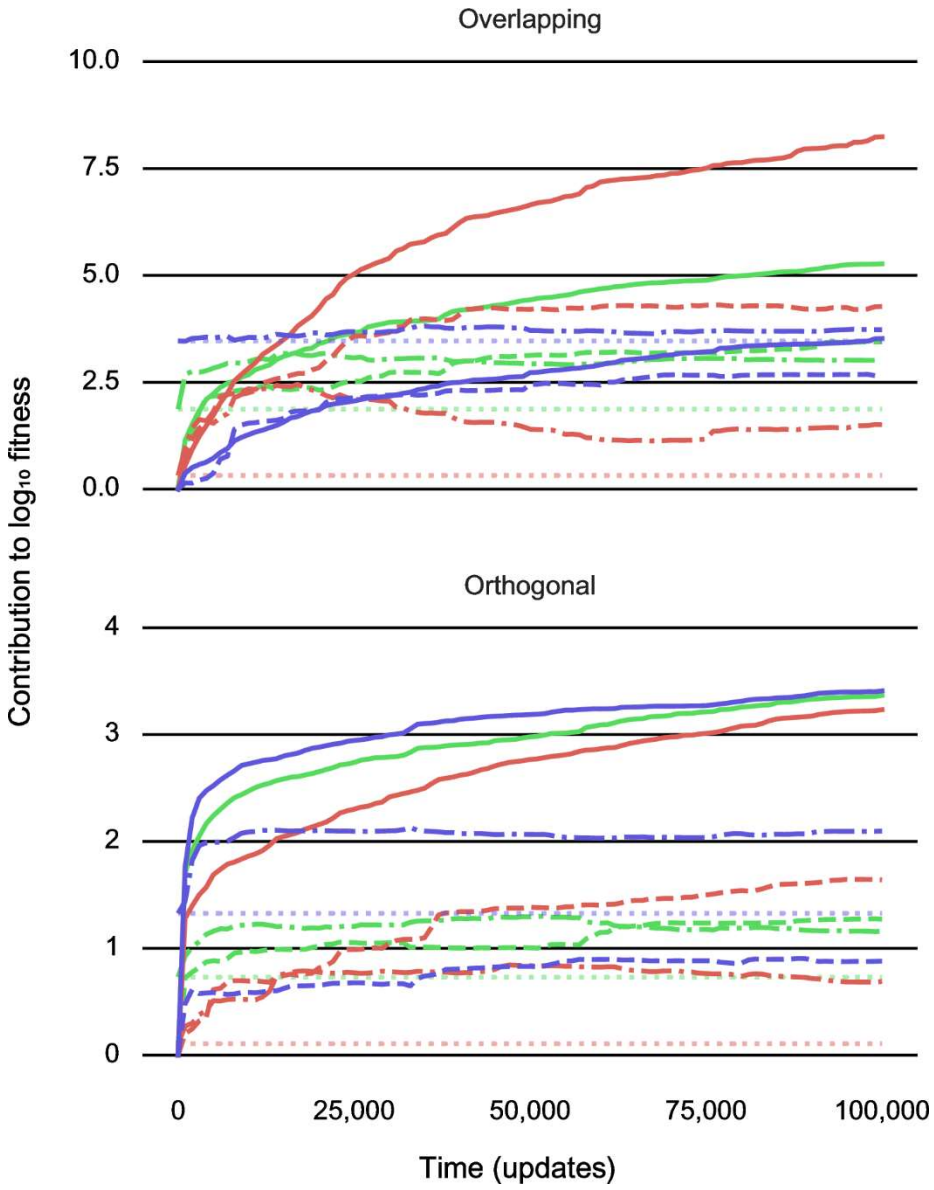


Figure 2.9. Trajectories for the contributions of adaptation, chance, and history to fitness in the new Phase II environment
Contributions of adaptation (solid lines), chance (dashed lines), and history (dash-dotted lines) to the evolution of fitness in the Overlapping (top) and Orthogonal (bottom) environments measured every 1,000 updates during Phase II for populations with shallow (red), intermediate (green), and deep (blue) histories from Phase I. The plot shows ancestral variances (dotted lines) for comparison.

Confidence intervals and confident inferences The results above are based on a total of 600 Phase II populations (2 environments x 3 depths of history x 100 replicates) and 2 focal traits (fitness and genome length). We did not emphasize statistical details when making comparisons,

as we highlighted differences that are generally well supported by our analyses. With respect to genome size (Figure 2.7), none of the six 95% confidence intervals for the contribution of adaptation overlap with the corresponding point estimate for either chance or history, nor do any of the confidence intervals for chance and history overlap the corresponding point estimate for adaptation. With respect to mean fitness, we saw an interesting interaction between the depth of history in Phase I and the environment during Phase II in terms of the contribution of adaptation to fitness in the new environment (Figure 2.8). To formalize that interaction, we calculated the following quantity for each Phase I progenitor and its 60 descendant populations: $Y = (A / B) - (C / D)$, where A is the mean change in $\log_{10}$-fitness in the Overlapping environment after Shallow history, B is the mean change in the Overlapping environment after Deep history, C is the mean change in the Orthogonal environment after Shallow history, and D is the mean change in the Orthogonal environment after Deep history. This difference in ratios thus expresses the environment-dependent change in the effect of the depth of history on the contribution of adaptation to fitness in the new environment. Under the null hypothesis of no interaction between the depth of history during Phase I and the environment in Phase II, the expected value of Y is 0. Each difference is independent of the others, and all 10 differences are positive, indicating a consistently greater difference in the contribution of adaptation between populations with shallow and deep histories in the Overlapping environment than in the Orthogonal environment. A non-parametric sign test confirms that this interaction is highly significant (two-tailed $p = 0.0020$).

We also performed a meta-analysis of the overall patterns and their consistency with respect to the effects of the depth of the historical footprint. Figures 2.7 and 2.8 combined include 16 cases where we estimated four contributions (adaptation, chance, history, and ancestral) for two traits at three different depths of history. If the effects of the depth of history were statistically

meaningless (i.e., if the depth of history's influence was random), then one would expect the estimated contribution for the intermediate depth to fall between the contributions for the shallow and deep histories only one-third of the time. In fact, the contribution at the intermediate depth was between the other two contributions in 14 of 16 cases. The probability of an outcome as or more extreme than that under the binomial distribution with an expected success rate of 1/3 is extremely low ($p \approx 0.00001$). Even if we exclude the four ancestral cases (because these measurements are the history estimate at the start rather than the end of Phase II), the contributions associated with the intermediate depth fell between the two other contributions in 10 of 12 cases, which remains highly significant ($p \approx 0.00054$). Only the effect of adaptation on genome length lacks a clear trend with respect to the depth of history, consistent with that trait having been selectively neutral or nearly so. These otherwise consistent outcomes reinforce the striking interaction between the depth of history and the reward structure of the new environment with respect to the contribution of adaptation to fitness (Figure 2.8).

Discussion

Overview

Gould famously proposed "replaying life's tape" to gain insight into the role of historical contingency in evolution [1]. Although he viewed this idea as a mere thought-experiment, others developed a quantitative framework for analyzing the contributions of adaptation, chance, and history to phenotypic evolution in experiments with replicated and nested lineages. And while Gould imagined replaying evolution from one moment in time, our experiments were designed to examine how the depth of history affected later evolution. In Phase I, we evolved 10 populations of digital organisms from a common ancestor in a single environment for periods representing shallow, intermediate, and deep histories. In Phase II, these 30 progenitors each founded 10 new

populations that evolved in two new environments. Using the statistical framework of Travisano et al. [3], we measured the contributions of adaptation, chance, and history to fitness and genome length. All three factors contributed significantly to the among-population variation in fitness (Figure 2.8). In the new environment with resources that overlapped those in the old environment, the deep footprint of history contributed as much as adaptation to fitness in the new environment. In the other new environment, which rewarded entirely different functions from the old environment, adaptation overcame even the deep footprint of history. By contrast, history and chance explained most of the variation in genome length, with almost no effect of adaptation (Figure 2.7). In the future, we hope other researchers will implement our experimental design for analyzing the effects of history's footprint using biological as well as digital organisms.

Perhaps our most noteworthy observation is that deepening the footprint of history constrained adaptation, but only when the old and new environments were similar. The Phase II Overlapping environment rewarded new functions while continuing to reward those functions that were useful in the ancestral environment. This overlap provided an evolutionary "bridge" that allowed lineages to retain adaptations that arose during Phase I. Deepening history by extending the duration of Phase I provided more time to acquire high-fitness phenotypes in the ancestral environment. These higher—and, importantly, also more variable—initial fitness levels were then sustained in the new Overlapping environment, providing a contribution of history to fitness that increased with the depth of history's footprint (Figure 2.8).

By contrast, the Orthogonal environment was largely "blind" to the phenotypic diversity that had been adaptive during Phase I. By chance, some Phase I lineages may have evolved the ability to perform functions that the Orthogonal environment would subsequently reward, even while they were still in the ancestral environment where these tasks were not rewarded. These pre-

adaptations could arise as correlated responses because an organism's ability to perform an unrewarded logic function was connected to its ability to perform a rewarded function. Populations with the deepest ancestry had the greatest opportunity to evolve these cryptic pre-adaptations. Thus, the ancestral variation in fitness at the beginning of Phase II in the Orthogonal environment increased as the footprint of history deepened, just as it did in the Overlapping environment (Figure 2.8). However, the ancestral variation—even after the deepest Phase I history—was rapidly overwhelmed by selection for new functions in the Orthogonal environment, where the contribution of adaptation was essentially independent of phylogenetic depth (Figure 2.9).

While adaptation was a potent force driving changes in fitness, that was not the case for genome length, where both chance and history contributed much more (Figure 2.7). Despite the differential contributions of adaptation, chance, and history to these two traits, it is noteworthy that deepening the footprint of history (by extending the duration of Phase I in the old environment) had consistent ordinal effects on the contributions of adaptation, chance, and history to both traits in each new environment. That is, the intermediate depth almost always yielded contributions between those of shallow and deep history (Figures 2.7 and 2.8), showing the importance of the depth of history's footprint on subsequent evolution. We also observed that adaptation, chance, and history were each capable of being the strongest driver of phenotypic evolution, depending on the specific circumstances of the trait, depth of history, and environment. Experimental approaches to replaying the tape of life thus offer a window into the "blind watchmaker's" [39] workshop, allowing one to examine how and when various forces contribute to the production of evolved form and function.

As noted by Blount et al. [2], the philosopher of science John Beatty distinguishes between two notions of contingency ascribed to Gould [40]. One is an unpredictability-centered definition

of "contingency *per se*" that identifies prior states that are *insufficient* to bring about a given outcome. This version seems to align with the randomness among replicates that we measured as chance. The other is a causal-dependency definition of "contingency upon", in which certain prior states are *necessary* to bring about a particular outcome. This version aligns more with the lineage effects that depend upon a particular progenitor, which we measured as history. Although these two notions are distinct, Beatty concludes they are compatible. He proposes that contingency might therefore be more fully understood as follows: "a historically contingent sequence of events is one in which the prior states are necessary or strongly necessary (causal-dependence version), but insufficient (unpredictability version) to bring about the outcome." Our approach and metrics are compatible with this integrated version, while our results show that deepening the footprint of history can shift the balance between these two aspects of contingency.

Future directions

Although we quantified the direction and magnitude of phenotypic changes in this work, we have not described them in terms of parallel and convergent evolution. These terms are widely used to describe similar changes, or homoplasy, in two or more independently evolving lineages. Russell Powell notes that many authors attribute homoplasy between closely related lineages to parallel evolution, whereas homoplasy between more distant lineages is attributed to selection's capacity to drive convergent outcomes [41,42]. Powell suggests using parallelism only when "a developmental homology is the proximate cause of the phenotypic similarity". He then describes a "screening off" test, adopted from Brandon [43], that could be used to distinguish between parallelism that results from internal, historically contingent constraints and convergence that is driven by external selection. The fine-grained data available with digital organisms should allow us to operationalize that test, or some variant of it, to determine the relative contributions of

parallelism and convergence to the homoplastic evolution of specific logic functions in our experiments. A future analysis could use our data to examine how the balance between parallelism and convergence changes with the depth of the footprint of history.

In this work, we also focused on evolutionary outcomes rather than the specific population-genetic processes that led to those outcomes. To examine such processes, Ostrowski et al. [44] used a two-phase experimental design to measure the contributions of antagonistic pleiotropy and mutation accumulation to niche specialization in digital organisms. Antagonistic pleiotropy occurs when a mutation that is beneficial overall nonetheless impairs some other function or functions; thus, it is fundamentally an adaptive process, albeit with tradeoffs. By contrast, mutation accumulation is a stochastic process that occurs when mutations that are not beneficial nonetheless increase by random drift or genetic linkage (e.g., hitchhiking with a beneficial mutation), leading to the reduction or loss of other functions. In the first phase of their experiment, Ostrowski et al. evolved generalists that used many different resources (performed many logic functions), which served as ancestors for populations that evolved in a single-resource environment during the second phase. They found that most mutations that caused the loss of previously rewarded functions were neutral or detrimental in the new environment, indicating the predominance of mutation accumulation. However, antagonistic pleiotropy also played an important role, as evidenced by an over-representation of mutations along the line of descent that were beneficial in the new environment but caused the loss of previously rewarded functions. These findings showed that both random and adaptive processes played important roles in ecological specialization, while also demonstrating that the contributions of those processes can be measured in Avida. Our experiment did not favor specialization, but instead involved a switch from one generalist environment to one of two other generalist environments: one that increased the number of

rewarded functions (Overlapping environment), and the other that rewarded a completely different set of functions (Orthogonal environment). Future analyses of our experiments might investigate the contributions of neutral and adaptive processes to the losses and gains of functions, to understand how their relative importance depends on the depth of history as well as on the selective environment. In the Orthogonal environment, we might expect antagonistic pleiotropy to become more important as the footprint of history deepens, because populations with a deep footprint are performing more unrewarded functions, so that more of their genome is a target for beneficial mutations that cause the loss of previously rewarded functions. In the Overlapping environment, by contrast, a shallow history means that more functions that are rewarded in both the old and new habitats remain available for discovery in Phase II, potentially generating synergistic pleiotropy. We expect synergistic pleiotropy would become less common as history deepens in the Overlapping environment, and it should be rare in the Orthogonal environment regardless of the depth of history.

Our study system has many advantages, as well as some potential limitations when making comparisons with biological systems. Avida was designed specifically to study evolution using digital organisms and it provides many tools that enabled us to observe the process in great detail. To that end, we have constructed and analyzed a self-contained evolving system, comparable in some ways to the *E. coli* Long-Term Evolution Experiment [17,45–54] and other microbial evolution experiments [3,55–61]. While it is difficult to compare evolving microbes to evolving computer programs, it is useful to recognize features of the digital world that are similar to the biological world, as well as others that are different. Some of the similarities include asexual reproduction (as in most but not all evolution experiments with microbes), the experimental tractability afforded by rapid generations, and the ability to store samples for later analysis. On the

other hand, the number of evolving populations in Avida is larger than is feasible in most microbial experiments, whereas the number of individuals per population is generally much larger when using microbes. Microbes also have much larger genomes, while per-site mutation rates are usually set much higher in Avida to offset the effects of smaller populations and genomes. The physiology and ecology of microorganisms and Avidians also undoubtedly differ profoundly. For these reasons, experiments with Avida are not meant to simulate biological evolution in any detail, but rather they offer a powerful way of testing general concepts and exploring questions in evolutionary science writ large. In any case, we hope others might use our experimental design—modified as needed for practicality—to study the effects of varying the depth of history's footprint on evolution in biological systems.

Concluding thoughts

In closing, we instantiated and extended Gould's thought-experiment [1] by "replaying life's tape" from multiple points in the history of these replicate virtual worlds. More than 40 years after Gould and Lewontin's famous critique of the "adaptationist programme" [62,63], our results also illustrate some of the complex ways in which adaptation, chance, and history together contribute to the evolutionary process. Last but not least, we think that our experimental design can be employed to study the effects of the depth of history's footprint in biological as well as digital systems.

Materials and Methods

Avida system and experiments

We used Avida (version 2.11) to perform the experiments and output the data. Avida is an open-source platform for experiments with digital organisms; it is freely available at

https://github.com/devosoft/avida. Charles Ofria led the development of Avida, and it is maintained by the Digital Evolution Lab at Michigan State University.

We started the Phase I populations with a proto-ancestral clone with a genome length of 100 instructions. It was unable to perform any logic functions. Each genome in Avida is a circular series of computational instructions, with 26 possible instructions at each position. While executing its genomic program, a digital organism can copy its genome; if it performs one or more rewarded functions, it obtains additional resources that allow it to run its program at a faster rate. While copying a genome, mutations occur at random; these include both point mutations, which change one instruction to another, and indels, which insert or delete a single instruction. In both Phases I and II, we set the point-mutation rate to 0.007 per instruction copied and the rate of indels to 0.00005 per instruction copied, with indels also occurring at a rate of 0.1 per post-replication division. In both phases, we set the maximum population size to 3600. When a digital organism completed its replication, the new offspring was placed in a randomly chosen position in the population, and the organism that was previously at that position was eliminated. These random deaths thus introduce genetic drift, which along with mutations generate stochasticity in the evolutionary process. Organisms also died if they executed a total number of instructions that exceeded 20 times their genome length; such deaths typically occur if a mutation caused an organism's genomic program to cycle indefinitely without reproducing itself.

In Avida, an organism's fitness is calculated as its rate of energy acquisition divided by the cost of producing an offspring, where that cost is equal to the number of instructions it must execute to copy its genome and divide. Note that an organism's execution length is distinct from (and greater than) its genome length. The execution length represents the total number of instructions performed by a digital organism prior to reproduction, which involves some

instructions being executed many times. Execution of an organism's genomic program includes not only copying the genome but also performing Boolean logic operations that may increase the organism's rate of energy acquisition.

As in biology, many mutations are deleterious. A mutation might, for example, damage or destroy the copy loop that allows a digital organism to reproduce; or a mutation might disrupt the performance of a logic function that previously evolved, thus reducing the rate at which that organism acquires the energy that fuels the execution of its genomic program. Many other mutations are neutral, having no effect on an organism's phenotype. Yet other mutations can be beneficial in various ways. Some mutations may reduce the number of instructions needed to copy the genome, thereby making replication a bit more efficient. Others may enable the performance of one of the logic functions rewarded with additional energy. No single instruction is sufficient to perform any of the rewarded functions. However, a single mutation of just the right type may produce a new function in the right genetic background.

The user-defined environment determines whether performance of a particular function is rewarded; this situation is analogous to the choice of which resources are provided in an evolution experiment with bacteria or yeast. In Avida, the environments are defined by two- and three-input Boolean logic functions. If an organism performs a rewarded function while executing its genomic program, then it will receive a reward in the form of additional energy (in the form of CPU cycles) to execute those instructions. All else equal, this additional energy will allow it to reproduce faster, thereby providing an evolutionary advantage. An organism is rewarded for a given function if it outputs a 32-bit string that precisely matches the string that would be expected if it performed the associated bit-wise logic function on specific input strings.

The three environments used in our study reward different subsets of 9 two-input and 68 three-input Boolean logic functions. We excluded the most difficult two-input function (known as "equals" or EQU), leaving a total of 76 other functions. Following Wagenaar and Adami [33], we ordered this set of 76 functions and split it into two halves, called "even" and "odd", in which adjacent pairs of functions tend to be of similar difficulty given the genomic instructions available in Avida [33]. The Phase I environment rewarded the 38 even functions. In Phase II, the "Overlapping" environment rewarded the same 38 even functions along with the 38 odd functions. The Orthogonal environment, by contrast, rewarded only the 38 odd functions, which were not rewarded during Phase I. The rewards for an organism that can perform two or more logic functions are multiplicative. The specific multipliers for each function, along with other experimental settings, are provided in the configuration and README files at the github repository for this paper: https:// https://github.com/bundyjay/Avida_footprint_of_history.

Analysis

In nested experimental designs like the one that we employed, the contributions of chance and history to the evolution of a trait are estimated using variance components [3,33]. The contribution of chance reflects the variance among Phase II replicate populations derived from the same Phase I progenitor, whereas the contribution of history reflects the additional variation (above that due to chance) associated with the different Phase I progenitors. Both chance and history reflect random, as opposed to fixed, sources of variation. We then constructed 95% confidence intervals around the point estimates using the chi-square distribution [64]. We transformed the variance estimates and their confidence limits by taking the square root of the values, so that the scaling is directly comparable to adaptation [3,33]. We estimated the contribution of adaptation as the mean difference in trait values between each set of 100 Phase II populations and their corresponding

progenitors [3,33]. We constructed 95% confidence intervals using the *t* distribution with 99 degrees of freedom (based on the number of Phase II populations). We measured the contributions of adaptation, chance, and history separately for the sets of populations founded from progenitors sampled at three points during Phase I (20,000, 100,000, and 500,000 updates), along with their confidence limits. This approach allowed us to examine how the footprint of history from Phase I affected evolution in both Phase II environments (Figures 2.7 and 2.8). We also measured the contributions of adaptation, chance, and history throughout Phase II, and we show these complete trajectories without their associated confidence limits for visual clarity (Figure 2.9 and 2.17).

Owing to the several order-of-magnitude changes in fitness over time, we $\log_{10}$-transformed fitness values prior to all analyses. For consistency, we similarly transformed genome lengths, although they varied much less. We performed all data analyses using Python (version 3).

Acknowledgments

We dedicate this chapter to the memory of George Gilchrist, who was the NSF program officer for the BEACON Center for the Study of Evolution in Action. George's dedication, scholarship, and advice are greatly missed. We thank Louise Mead and Arend Hintze for valuable discussions, and members of the Ofria and Lenski labs for their collegial input. This research received partial support from the BEACON Center (NSF cooperative agreement DBI-0939454) along with fellowships and the John Hannah endowment from Michigan State University.

APPENDIX

APPENDIX

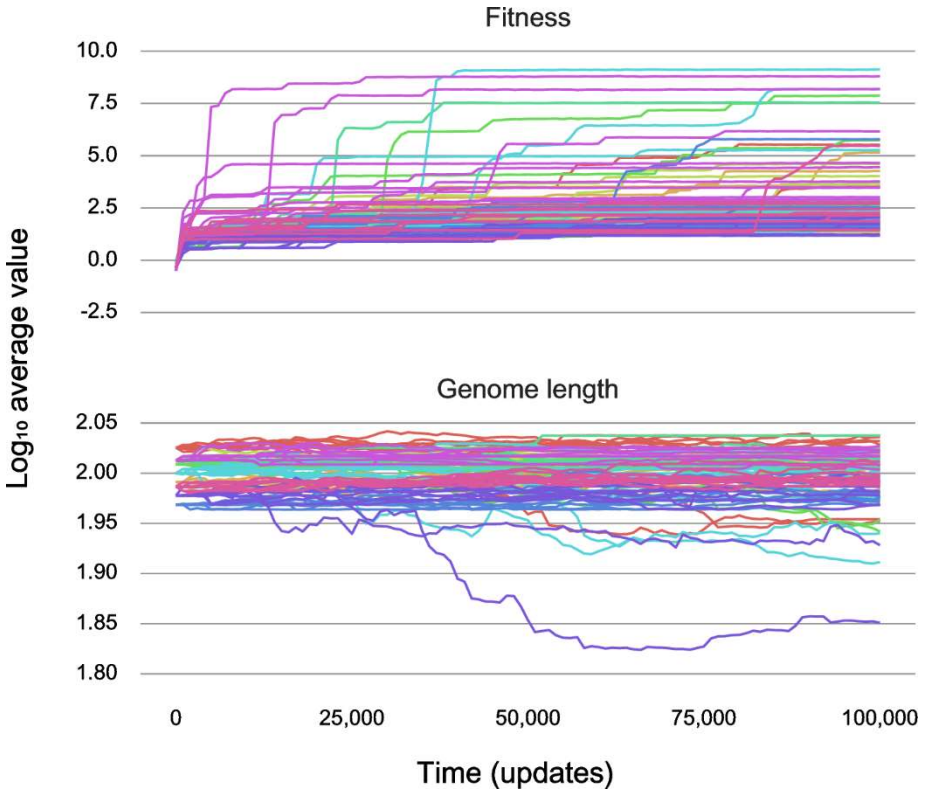


Figure 2.10. Population averages for fitness and genome length during Phase II in the Orthogonal environment for populations with a shallow history
Average fitness (top) and genome length (bottom) measured every 1,000 updates for 10 replicate populations derived from a single founder from each Phase I population, as they evolved for 100,000 updates in the new Phase II environment. These populations have a shallow history, as their founders evolved for only 20,000 updates in the ancestral environment. The new environment rewarded performance of a set of 38 logic functions that do not overlap with those rewarded in the ancestral environment.

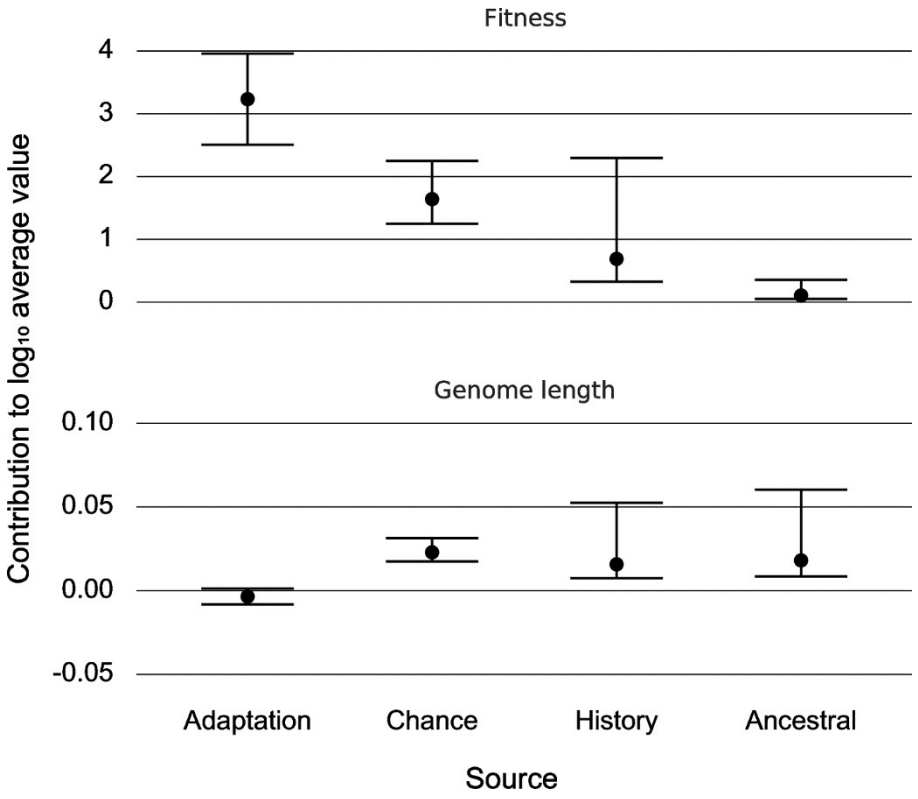


Figure 2.11. End-of-experiment effects of adaptation, chance, and history in the Orthogonal environment with shallow history

Contributions of adaptation, chance, and history to the evolution of fitness (top) and genome length (bottom) based on end-of-experiment average values for populations with a shallow history in Phase I that evolved in the Orthogonal environment during Phase II. Error bars show 95% confidence limits for the associated contributions, along with the ancestral variation at the start of Phase II.

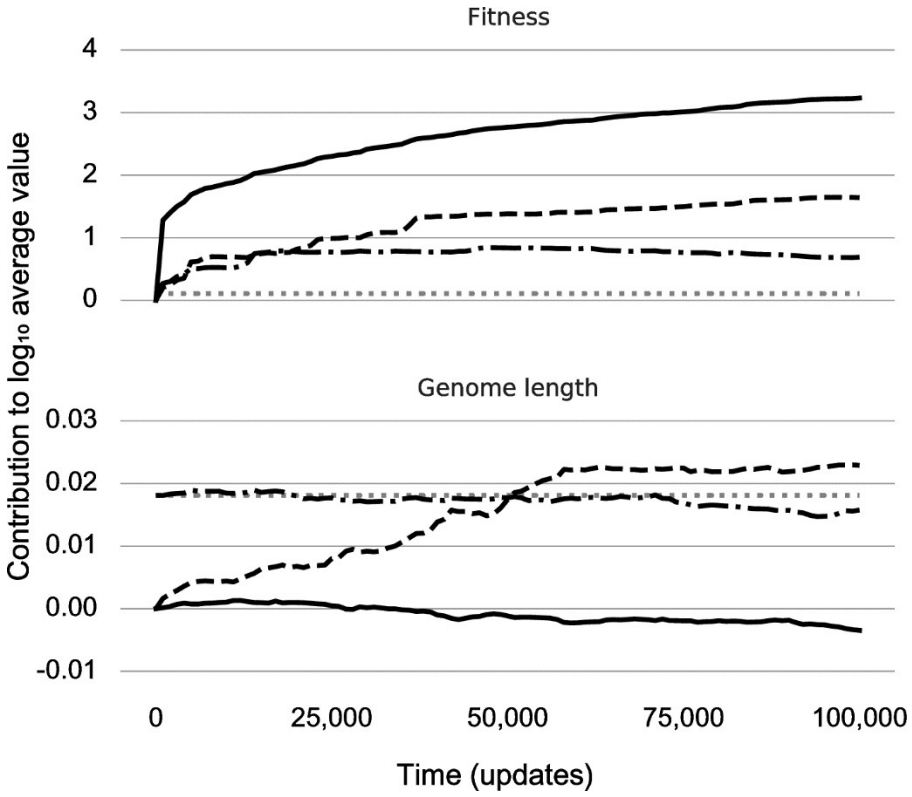


Figure 2.12. Trajectories of adaptation, chance, and history in the Orthogonal environment with shallow history

Contributions of adaptation (solid line), chance (dashed line), and history (dash-dotted line) to the evolution of fitness (top) and genome length (bottom) measured every 1,000 updates during Phase II. This plot shows the ancestral variance (dotted line) for comparison.

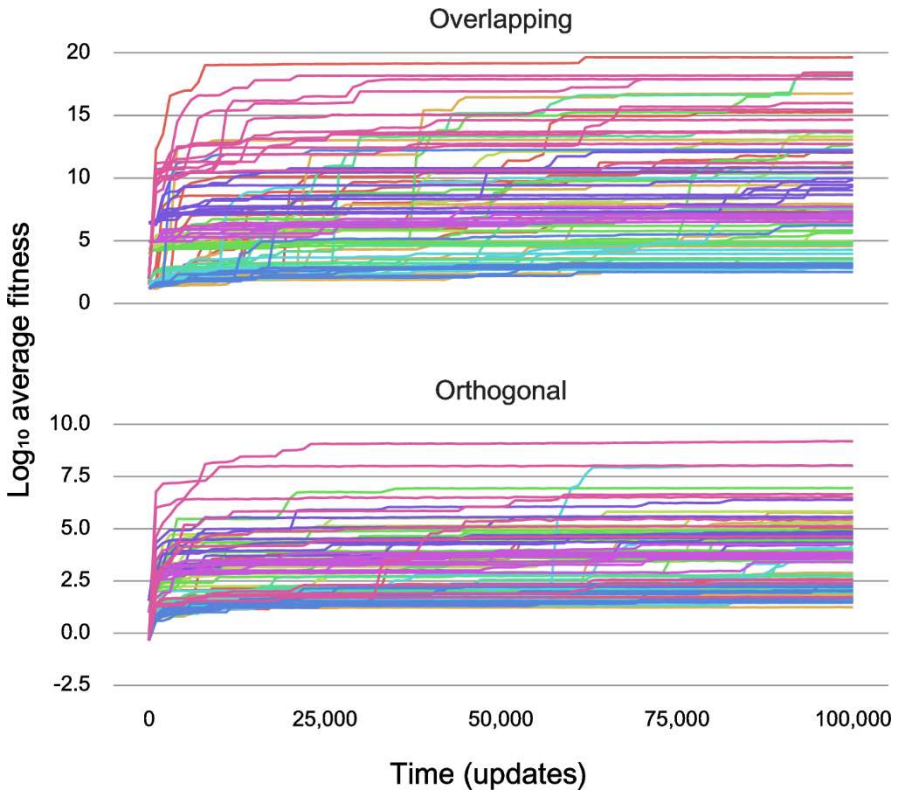


Figure 2.13. Population averages for fitness during Phase II with intermediate history
Average fitness in the Overlapping (top) and Orthogonal (bottom) environments measured every 1,000 updates for 10 replicate populations derived from a single founder from each Phase I population, as they evolved for 100,000 updates in the new Phase II environment. These populations have an intermediate history, as their founders evolved for 100,000 updates in the ancestral environment.

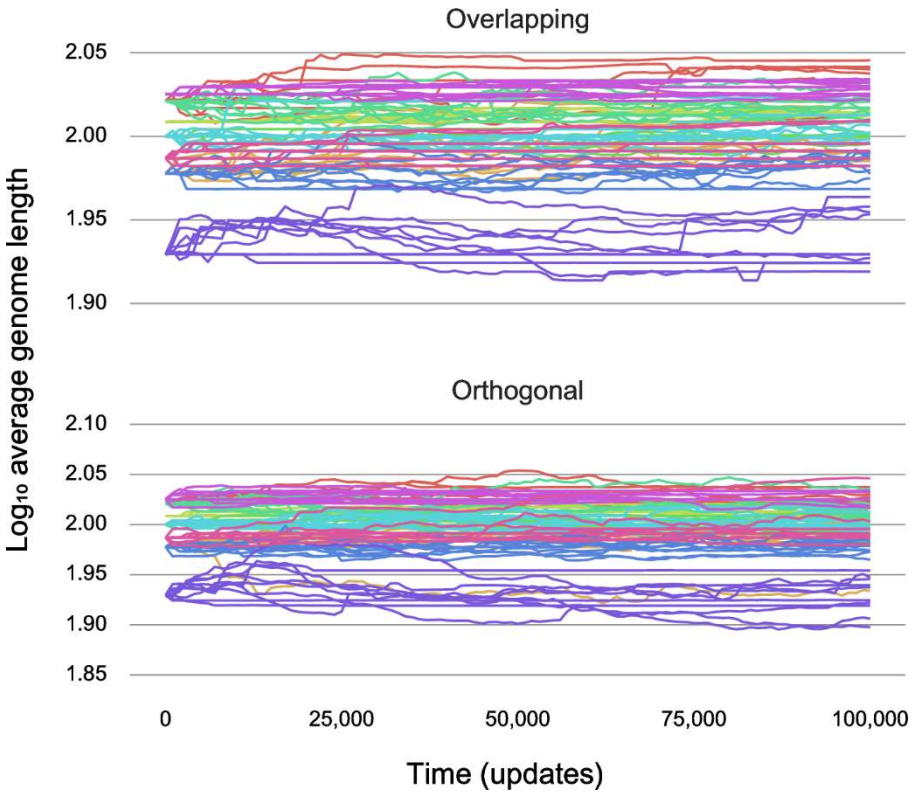


Figure 2.14. Population averages for genome length during Phase II with intermediate history

Average genome length in the Overlapping (top) and Orthogonal (bottom) environments measured every 1,000 updates for 10 replicate populations derived from a single founder from each Phase I population, as they evolved for 100,000 updates in the new Phase II environment. These populations have an intermediate history, as their founders evolved for 100,000 updates in the ancestral environment.

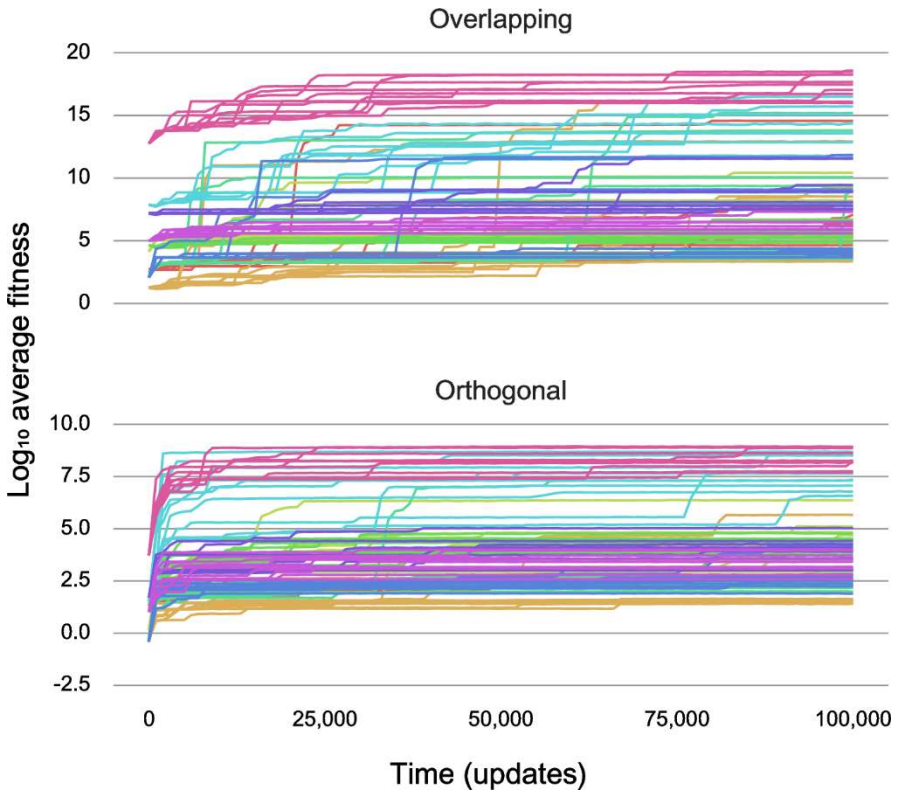


Figure 2.15. Population averages for fitness during Phase II with deep history
Average fitness in the Overlapping (top) and Orthogonal (bottom) environments measured every 1,000 updates for 10 replicate populations derived from a single founder from each Phase I population, as they evolved for 100,000 updates in the new Phase II environment. These populations have a deep history, as their founders evolved for 500,000 updates in the ancestral environment.

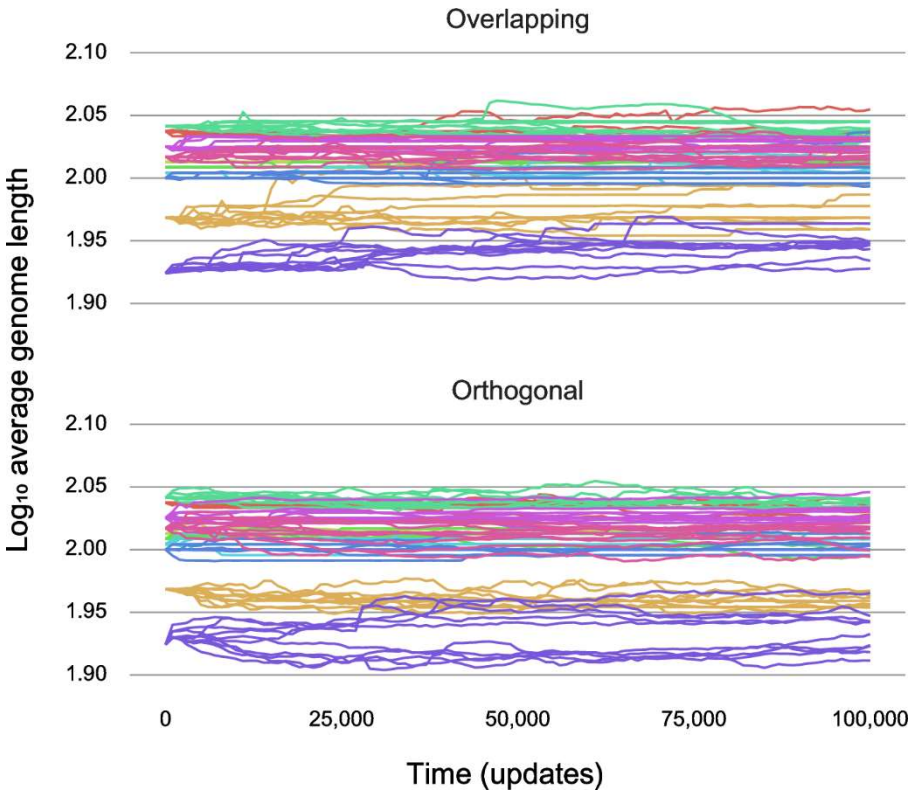


Figure 2.16. Population averages for genome length during Phase II with deep history
Average genome length in the Overlapping (top) and Orthogonal (bottom) environments measured every 1,000 updates for 10 replicate populations derived from a single founder from each Phase I population, as they evolved for 100,000 updates in the new Phase II environment. These populations have a deep history, as their founders evolved for 500,000 updates in the ancestral environment.

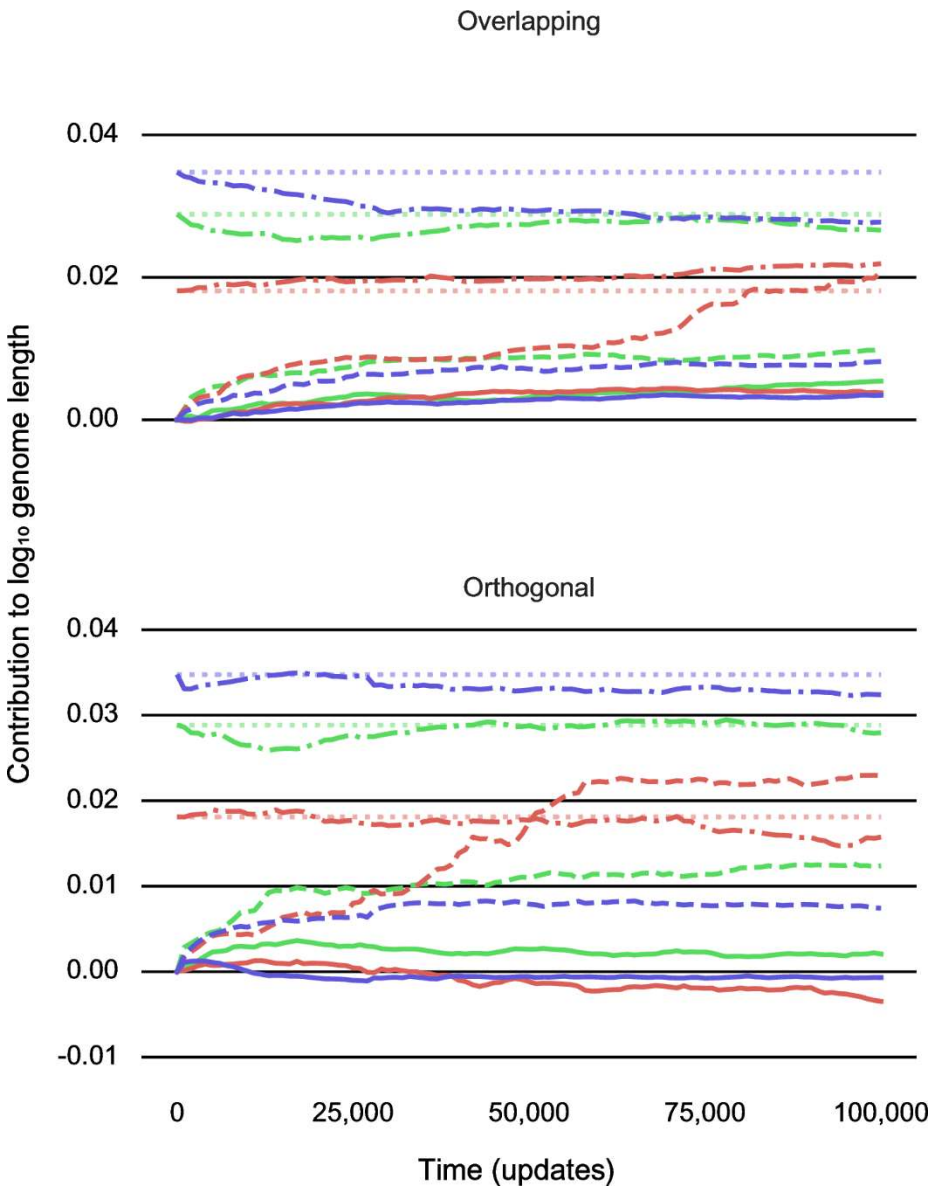


Figure 2.17. Trajectories for the contributions of adaptation, chance, and history to genome length during Phase II

Contributions of adaptation (solid lines), chance (dashed lines), and history (dash-dotted lines) to the evolution of genome length in the Overlapping (top) and Orthogonal (bottom) environments measured every 1,000 updates during Phase II for populations with shallow (red), intermediate (green), and deep (blue) Phase I histories. This plot shows the ancestral variances (dotted lines) for comparison.

REFERENCES

REFERENCES

1. Gould SJ. Wonderful Life. W. W. Norton; 1989.

2. Blount ZD, Lenski RE, Losos JB. Contingency and determinism in evolution: Replaying life's tape. Science. 2018; 362(6415):1–10.

3. Travisano M, Mongold JA, Bennett AF, Lenski RE. Experimental tests of the roles of adaptation, chance, and history in evolution. Science. 1995; 267(5194):87–90.

4. Losos JB, Jackman TR, Larson A, de Queiroz K, Rodriguez-Schettino L. Contingency and determinism in replicated adaptive radiations of island lizards. Science. 1998; 279(5359):2115–2118.

5. Teotónio H, Rose MR. Variation in the reversibility of evolution. Nature. 2000; 408(6811):463–466.

6. Emerson SB. A macroevolutionary study of historical contingency in the fanged frogs of Southeast Asia. Biol J Linn Soc Lond. 2001; 73(1):139–151.

7. Vanhooydonck B, Irschick DJ. Is evolution predictable? Evolutionary relationships of divergence in ecology, performance and morphology in Old and New World lizard radiations. Topics in Functional and Ecological Vertebrate Morphology. 2002; 191–204.

8. Morris SC. Life's Solution. Cambridge Univ Press; 2003.

9. Joshi A, Castillo RB, Mueller LD. The contribution of ancestry, chance, and past and ongoing selection to adaptive evolution. J Genet. 2003; 82(3):147–162.

10. Ortlund EA, Bridgham JT, Redinbo MR, Thornton JW. Crystal structure of an ancient protein: evolution by conformational epistasis. Science. 2007; 317(5844):1544–1548.

11. Keller SR, Taylor DR. History, chance and adaptation during biological invasion: Separating stochastic phenotypic evolution from response to selection. Ecol Lett. 2008; 11(8):852–866.

12. Dick MH, Lidgard S, Gordon DP, Mawatari SF. The origin of *Ascophoran bryozoans* was historically contingent but likely. Proc R Soc B. 2009; 276(1670):3141–3148.

13. Conway Morris S. Evolution: like any other science it is predictable. Phil Trans R Soc B. 2010; 365(1537):133–145.

14. Meyer JR, Dobias DT, Weitz JS, Barrick JE, Quick RT, Lenski RE. Repeatability and contingency in the evolution of a key innovation in phage lambda. Science. 2012; 335(6067):428–432.

15. Harms MJ, Thornton JW. Historical contingency and its biophysical basis in glucocorticoid receptor evolution. Nature. 2014; 512(7513):203–207.

16. Starr TN, Picton LK, Thornton JW. Alternative evolutionary histories in the sequence space of an ancient protein. Nature. 2017; 549(7672):409–413.

17. Lenski RE, Rose MR, Simpson SC, Tadler SC. Long-term experimental evolution in *Escherichia coli*. I. Adaptation and divergence during 2,000 generations. Am Nat. 1991; 138(6):1315–1341.

18. Flores-Moya A, Costas E, López-Rodas V. Roles of adaptation, chance and history in the evolution of the dinoflagellate *Prorocentrum triestinum*. Naturwissenschaften. 2008; 95(8):697–703.

19. Flores-Moya A, Rouco M, García-Sánchez MJ, García-Balboa C, González R, Costas E, et al. Effects of adaptation, chance, and history on the evolution of the toxic dinoflagellate *Alexandrium minutum* under selection of increased temperature and acidification. Ecol Evol. 2012; 2(6):1251–1259.

20. Pérez-Zaballos FJ, Ortega-Mora LM, Álvarez-García G, Collantes-Fernández E, Navarro-Lozano V, García-Villada L, et al. Adaptation of *Neospora caninum* isolates to cell-culture changes: an argument in favor of its clonal population structure. J Parasitol. 2005; 507–510.

21. Rouco M, López-Rodas V, Flores-Moya A, Costas E. Evolutionary changes in growth rate and toxin production in the cyanobacterium *Microcystis aeruginosa* under a scenario of eutrophication and temperature increase. Microb Ecol. 2011; 62(2):265–273.

22. Rebolleda-Gómez M, Travisano M. Adaptation, chance, and history in experimental evolution reversals to unicellularity. Evolution. 2019; 73(1):73–83.

23. Santos-Lopez A, Marshall CW, Welp AL, Turner C, Rasero J, Cooper VS. The roles of history, chance, and natural selection in the evolution of antibiotic resistance. bioRxiv. 2020; 2020.07.22.216465.

24. Burch CL, Lin C. Evolvability of an RNA virus is determined by its mutational neighbourhood. Nature. 2000; 406(6796):625–628.

25. Teotónio H, Chelo IM, Bradić M, Rose MR, Long AD. Experimental evolution reveals natural selection on standing genetic variation. Nat Genet. 2009; 41(2):251–257.

26. Bedhomme S, Lafforgue G, Elena SF. Genotypic but not phenotypic historical contingency revealed by viral experimental evolution. BMC Evol Biol. 2013; 13(1):46.

27. Kryazhimskiy S, Rice DP, Jerison ER, Desai MM. Global epistasis makes adaptation predictable despite sequence-level stochasticity. Science. 2014; 344(6191):1519–1522.

28. Spor A, Kvitek DJ, Nidelet T, Martin J, Legrand J, Dillmann C, et al. Phenotypic and genotypic convergences are influenced by historical contingency and environment in yeast. Evolution. 2014; 68(3):772–790.

29. Simões P, Fragata I, Seabra SG, Faria GS, Santos MA, Rose MR, et al. Predictable phenotypic, but not karyotypic, evolution of populations with contrasting initial history. Sci Reports. 2017; 7(1):913.

30. Taylor T, Hallam J. Replaying the tape: An investigation into the role of contingency in evolution. In: Artificial Life VI: Proceedings of the Sixth International Conference on Artificial Life. MIT Press; 1998. p. 256–265.

31. Yedid G, Bell G. Macroevolution simulated with autonomously replicating computer programs. Nature. 2002; 420(6917):810–812.

32. Lenski RE, Ofria C, Pennock RT, Adami C. The evolutionary origin of complex features. Nature. 2003; 423(6936):139–144.

33. Wagenaar DA, Adami C. Influence of chance, history, and adaptation on digital evolution. Artificial Life. 2004; 10(2):181–190.

34. Braught G, Dean A. The effects of learning on the roles of chance, history and adaptation in evolving neural networks. In: Australian Conference on Artificial Life. Berlin, Heidelberg: Springer; 2007; p. 201–211.

35. Yedid G, Ofria CA, Lenski RE. Historical and contingent factors affect re-evolution of a complex feature lost during mass extinction in communities of digital organisms. J Evol Biol. 2008; 21(5):1335–1357.

36. Adami C. Introduction to Artificial Life. Springer; 1998.

37. Lenski RE, Ofria C, Collier TC, Adami C. Genome complexity, robustness and genetic interactions in digital organisms. Nature. 1999; 400(6745):661–664.

38. Wilke CO, Wang JL, Ofria C, Lenski RE, Adami C. Evolution of digital organisms at high mutation rates leads to survival of the flattest. Nature. 2001; 412(6844):331–333.

39. Dawkins R. The Blind Watchmaker. W. W. Norton; 2015.

40. Beatty J. Replaying life's tape. J Philosophy. 2006; 103(7):336–362.

41. Powell R. Is convergence more than an analogy? Homoplasy and its implications for macroevolutionary predictability. Biol Philos. 2007; 22(4):565–578.

42. Powell R. Contingency and convergence in macroevolution: A reply to John Beatty. J Philosophy. 2009; 106(7):390–403.

43. Brandon RN. Adaptation and Environment. Princeton Univ Press; 1990.

44. Ostrowski EA, Ofria C, Lenski RE. Ecological specialization and adaptive decay in digital organisms. Am Nat. 2007; 169(1):E1–20.

45. Lenski RE, Travisano M. Dynamics of adaptation and diversification: A 10,000-generation experiment with bacterial populations. Proc Natl Acad Sci USA. 1994; 91(15):6808–6814.

46. Sniegowski PD, Gerrish PJ, Lenski RE. Evolution of high mutation rates in experimental populations of *E. coli*. Nature. 1997; 387(6634):703–705.

47. Cooper VS, Lenski RE. The population genetics of ecological specialization in evolving *Escherichia coli* populations. Nature. 2000; 407(6805):736–739.

48. Blount ZD, Borland CZ, Lenski RE. Historical contingency and the evolution of a key innovation in an experimental population of *Escherichia coli*. Proc Natl Acad Sci USA. 2008; 105(23):7899–7906.

49. Khan AI, Dinh DM, Schneider D, Lenski RE, Cooper TF. Negative epistasis between beneficial mutations in an evolving bacterial population. Science. 2011; 332(6034):1193–1196.

50. Blount ZD, Barrick JE, Davidson CJ, Lenski RE. Genomic analysis of a key innovation in an experimental *Escherichia coli* population. Nature. 2012; 489(7417):513–518.

51. Good BH, McDonald MJ, Barrick JE, Lenski RE, Desai MM. The dynamics of molecular evolution over 60,000 generations. Nature. 2017; 551(7678):45–50.

52. Card KJ, LaBar T, Gomez JB, Lenski RE. Historical contingency in the evolution of antibiotic resistance after decades of relaxed selection. PLoS Biol. 2019; 17(10).

53. Blount ZD, Maddamsetti R, Grant NA, Ahmed ST, Jagdish T, Baxter JA, et al. Genomic and phenotypic evolution of *Escherichia coli* in a novel citrate-only resource environment. Elife. 2020; 9:e55414.

54. Grant NA, Maddamsetti R, Lenski RE. Maintenance of metabolic plasticity despite relaxed selection in a long-term evolution experiment with *Escherichia coli*. Am Nat. 2021; in press.

55. Rainey PB, Travisano M. Adaptive radiation in a heterogeneous environment. Nature. 1998; 394(6688):69–72.

56. Kawecki TJ, Lenski RE, Ebert D, Hollis B, Olivieri I, Whitlock MC. Experimental evolution. Trends Ecol Evol. 2012; 27(10):547–560.

57. Ratcliff WC, Denison RF, Borrello M, Travisano M. Experimental evolution of multicellularity. Proc Natl Acad Sci USA. 2012; 109(5):1595–1600.

58. Barrick JE, Lenski RE. Genome dynamics during experimental evolution. Nat Rev Gen. 2013; 14(12):827–839.

59. Lang GI, Rice DP, Hickman MJ, Sodergren E, Weinstock GM, Botstein D, et al. Pervasive genetic hitchhiking and clonal interference in forty evolving yeast populations. Nature. 2013; 500(7464):571–574.

60. Levy SF, Blundell JR, Venkataram S, Petrov DA, Fisher DS, Sherlock G. Quantitative evolutionary dynamics using high-resolution lineage tracking. Nature. 2015; 519(7542):181–186.

61. Johnson MS, Gopalakrishnan S, Goyal J, Dillingham ME, Bakerlee CW, Humphrey PT, et al. Phenotypic and molecular evolution across 10,000 generations in laboratory budding yeast populations. Elife. 2021; 10:e63910.

62. Gould SJ, Lewontin RC. The spandrels of San Marco and the Panglossian paradigm: A critique of the adaptationist programme. Proc R Soc B. 1979; 205(1161):581–598.

63. Price RM, Perez KE. Beyond the adaptationist legacy: Updating our teaching to include a diversity of evolutionary mechanisms. Am Biol Teacher. 2016; 78(2):101–108.

64. Sokal RR, Rohlf J. Biometry. 3rd ed. New York, NY: W. H. Freeman; 1994.

CHAPTER 3: HABITAT NOVELTY AND HISTORY SHAPE TRADE-OFFS AND TRADE-UPS IN DIGITAL ORGANISMS

Authors: Jason Nyerere Bundy, Charles Ofria, and Richard E. Lenski

Abstract

Understanding how new traits evolve under changing conditions is an important issue in biology. Few experiments have investigated how historical divergence between founding genotypes influences genomic constraints on lineages evolving in and adapting to new environments. We analyzed a two-phase experiment to determine how the interaction of historical effects and habitat novelty influenced the nature of pleiotropy associated with gain-of-function mutations in evolving populations of digital organisms. In the case of a new environment that shares many features of the ancestral habitat, we saw that populations founded by ancestors with shallow history (i.e., recently diverged) were more likely to evolve new functions through mutations that are beneficial in both the new and ancestral habitats (i.e., trade-ups or synergistic pleiotropy). By contrast, populations with deeper historical divergence suffered frequent trade-offs (i.e., antagonistic pleiotropy) between fitness in the new and ancestral environments. The trade-offs associated with deep history in this environment help to explain why the populations derived from more primitive ancestors (i.e., shallow history) adapted more rapidly to the new environment. However, when the ancestral and novel environments held entirely different resources, trade-offs were ubiquitous, regardless of whether the founders had a shallow or deep history. In effect, the different selective forces erased the constraints of deep history imposed by previously evolved traits that were no longer adaptive, allowing lineages with shallow and deep history to adapt at similar rates. In sum, our experiments show that the depth of history can affect subsequent adaptation to new environments, at least in part, by altering the form of pleiotropy. Moreover, the extent of change between the old and new habitats can alter the effects of deep history.

Introduction

The combination of natural selection, chance processes (such as mutation and genetic drift), and the idiosyncratic effects of historical contingency drive evolutionary change. Selection can produce parallel and convergent outcomes, i.e., the repeated evolution of similar traits in closely and distantly related lineages, respectively [1]. By contrast, contingency and chance promote unpredictable outcomes [1]. To date, few experiments have examined how the extent of historical divergence between lineages shapes the genetic constraints and opportunities that can promote either convergent or contingent outcomes for descendant lineages evolving in new environments [2,3]. It seems plausible, for example, that lineages with a deeper history in the ancestral environment will face more antagonistic pleiotropy than lineages derived from more recently diverged founders, because mutations that increased specialization in the ancestral environment may impose trade-offs in the new environment. These trade-offs can restrict the niche breadth of descendants and thereby limit the availability of genetic paths leading to high fitness in the new environment [2,4–7]. Trade-offs are critical to understanding many ecological and evolutionary processes including community assembly [8–10], aging [11], life-history strategies [12–19], specialization [20–22], and adaptive radiations [23]. However, synergistic pleiotropy—that is, mutations with positively correlated fitness effects on multiple traits or across environments—also sometimes plays an important role in adaptive evolution [24–34]. Synergistic pleiotropy of beneficial mutations across new and ancestral habitats, sometimes called “trade-ups”, may facilitate the exploration of more and broader genetic paths to adaptation [2,22,30]. However, generalist phenotypes may not perform as well as more specialized types in any particular environment [30,35–40]. Thus, both antagonistic and synergistic pleiotropy can influence the propensity for later innovation [30].

A fundamental question in evolutionary biology is how the historical divergence of lineages affects the later evolution of novel functions and behaviors, especially when environments change [41–51]. Since Gould's thought experiment in *Wonderful Life* [52], several studies have turned his famous metaphor of "replaying life's tape" into an experimental program to examine contingency and convergence. These "replay" experiments seek to determine how variation between lineages that evolved from a common ancestor under identical conditions shapes the contributions of adaptation, chance, and history to subsequent evolution [3,53–62]. These studies show that closely related lineages often respond to changed environments in similar ways [1]. But researchers have yet to address fully two other related questions. First, how does deeper history and the resulting greater divergence between ancestors affect the potential for future contingency and convergence? Second, what circumstances (e.g., the extent of environmental change) promote more or less similar outcomes, even for distantly related lineages? We performed a two-phase evolution experiment with digital organisms—computer programs that self-replicate, mutate, compete, and evolve in virtual environments [63–65]—that can address both of these questions in the same system [66].

We previously analyzed this experiment to quantify the contributions of adaptation, chance, and history to fitness and genome length following a change in environment, including how these contributions depended on the depth of history in the ancestral environment and the extent of difference between the ancestral and new environments (Figure 3.1). In this paper, we examine how these same two factors—the depth of history, and the extent of environmental change—impact the pleiotropic effects of gain-of-function mutations that arose while the lineages evolved in and adapted to the new environments. We show that the interaction of these factors substantially influences the predominant form of pleiotropy (i.e., antagonistic or synergistic).

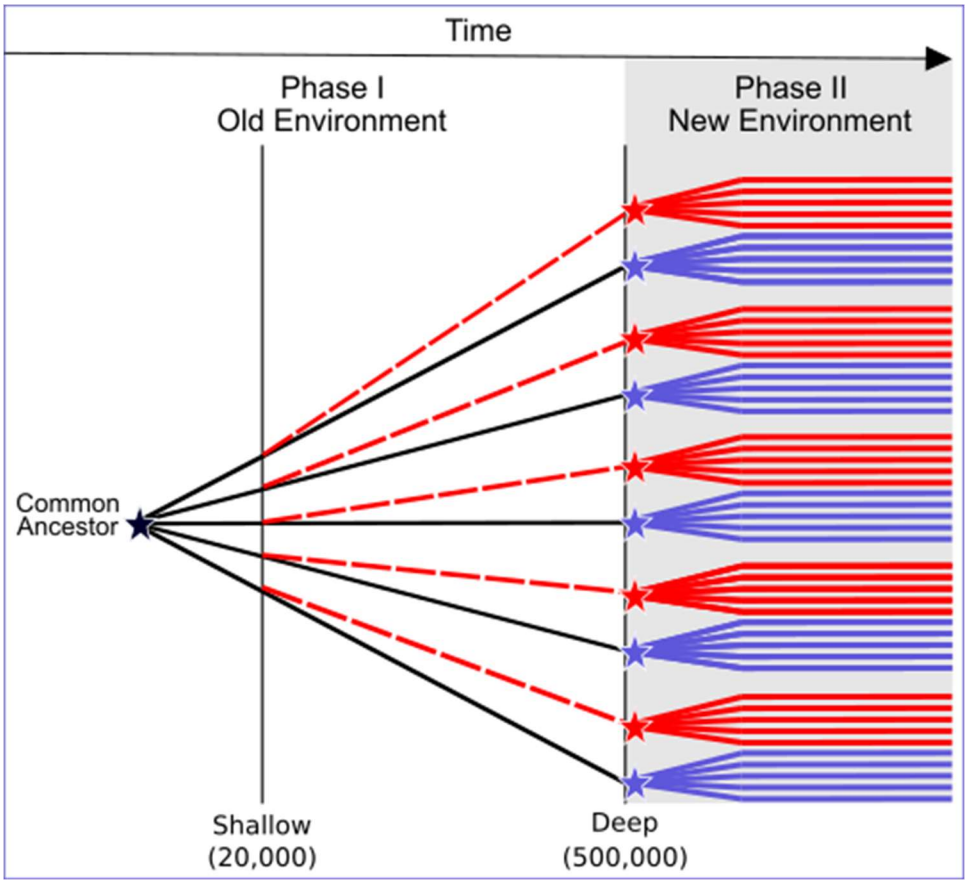


Figure 3.1. Schematic illustration of the experimental design
In Phase I, ten populations (five depicted) evolved from a single proto-ancestor (black star) for 500,000 updates (the unit of time in Avida). We sampled the most abundant genotype (colored stars) from each population after 20,000 (red) and 500,000 (blue) updates, representing shallow and deep history, respectively. In Phase II, each of the 20 proximate ancestors founded ten replicate populations (five depicted) in each of two new environments (one depicted), where they then evolved for 100,000 updates. One new environment (Overlapping) shared many resources with the old environment used in Phase I. The other new environment (Orthogonal) had a completely novel set of resources. Modified from [66].

Reviewing the experimental design, In Phase I we evolved 10 replicate populations from a common ancestor under identical conditions. These populations served as simultaneous, rather than sequential, replays of evolution, similar to the design of Lenski's *E. coli* Long-term Evolution Experiment [61,67,68]. Although the populations were identical at the start of Phase I, they eventually attained fitness values that varied over several orders of magnitude (Figure 3.2) [66]. We then isolated the numerically dominant genotype from each Phase I population at two different

timepoints, which served as the founders (proximate ancestors) for the Phase II populations. In Phase II, we evolved 10 replicate populations from each of the founders in two new environments, one similar (Overlapping) and one dissimilar (Orthogonal) to the old environment used in Phase I. The two timepoints at which we sampled the dominant genotypes from the Phase I populations correspond to shallow history (20,000 updates, or about 3,000 generations on average) and deep history (500,000 updates or about 65,000 generations on average).

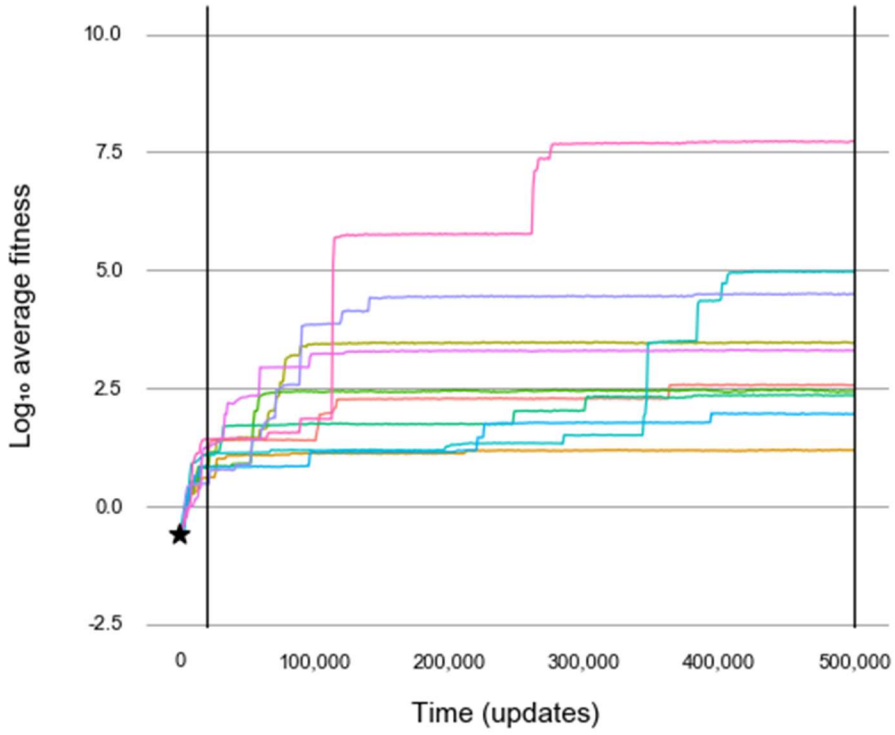


Figure 3.2. Fitness trajectories during Phase I
Log-transformed average fitness measured every 1,000 updates throughout Phase I for 10 replicate populations derived from a common ancestor (star). Vertical lines at 20,000 and 500,000 updates show the times when the dominant genotype was sampled from each population. These genotypes then served as the founders (proximate ancestors) for the Phase II populations.

In our previous analysis [66], we estimated the contributions of adaptation, chance, and history to the evolution of fitness at the end of Phase II using the analytical approach of Travisano et al. [53]. The estimate of adaptation reflects the change in the grand mean between the evolved populations and their founders from the start of Phase II. The estimate of chance reflects the

variation among replicate populations derived from the same founder (i.e., proximate ancestor). The contribution of history reflects the variation among the groups of replicates derived from different founders that we sampled at the same depth of history (shallow or deep).

We found that deep history constrained adaptation, but only when the Phase II populations evolved in the Overlapping environment that was more similar to the ancestral habitat (Figure 3.3). This experiment thus demonstrated a significant effect of deepening the "footprint of history" on adaptation. It also revealed conditions that mitigate that effect; in particular, the contribution of adaptation was essentially independent of the depth of history in the Orthogonal environment that was dissimilar to the ancestral environment. Therefore, this system is well-suited to studying how the depth of history and the extent of environmental change influence the form of pleiotropy for gain-of-function mutations that arose during Phase II.

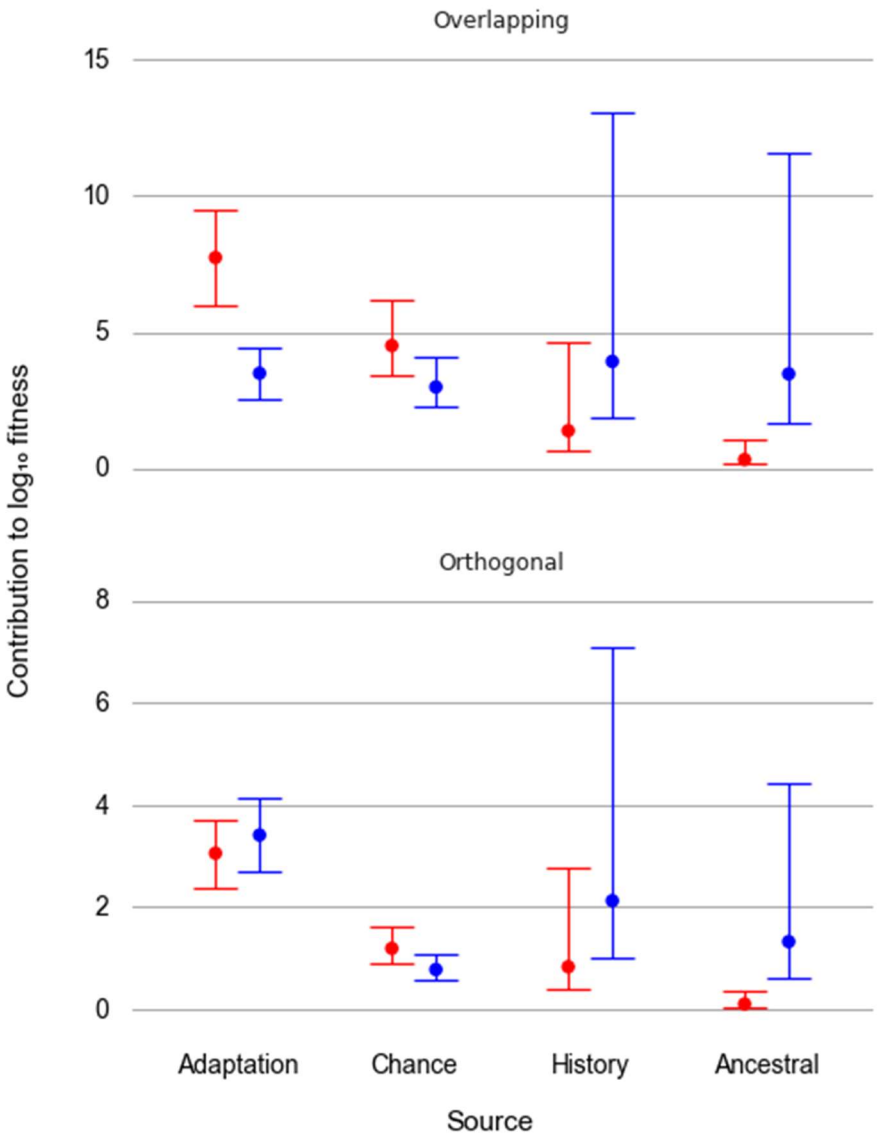


Figure 3.3. Impact of deep history on adaptation depends on extent of environmental change Contributions of adaptation, chance, and history to mean fitness for populations at the end of Phase II with shallow (red) and deep (blue) history in Phase I. Contributions are shown for the Overlapping (top) and Orthogonal (bottom) environments used during Phase II, which are similar and dissimilar to the Phase I environment, respectively. Error bars show 95% confidence limits. The graph also shows the ancestral variation at the start of Phase II for comparison. These results are quantitatively similar (and qualitatively identical) to a previous analysis of data obtained from an independent set of experiments run under identical conditions using an earlier version of the Avida program [66].

To that end, we analyzed how the degree of novelty imposed by the Phase II environment influenced the prevalence of antagonistic and synergistic pleiotropy for lineages with a shallow and deep history in the Phase I environment. We isolated every genotype along the line of descent leading to the numerically dominant genotype in each population at the end of Phase II. We identified all those mutations that resulted in the gain of new functions, and we determined each mutation's fitness effect in both the old (Phase I) and new (Phase II) environments. We used these fitness changes to characterize whether each mutation's pleiotropic effects resulted in antagonistic or synergistic pleiotropy. We also assessed the cumulative effect of all mutations along the line of descent on fitness in both the old and new environments to determine whether the net change led to a trade-off or a trade-up.

We wanted to test two specific hypotheses about how the interaction of the depth of history and the degree of environmental novelty influenced pleiotropy [66]. First, we hypothesized that populations with deep history would exhibit less *synergistic pleiotropy* than populations with a shallow history in the Overlapping environment, when conditions in the old and new environments were more similar (i.e., rewarded many of the same functions). We reasoned that more functions that had been rewarded in the old habitat would remain available at the outset of Phase II for lineages with shallow history, in comparison with deep history, owing to the shorter period for their ancestors to have already evolved those adaptations.

Second, we hypothesized that the populations with deep history would experience more *antagonistic pleiotropy* than populations with a shallow history in the Orthogonal environment, which rewarded an entirely different set of functions than did the old environment. The ancestors of the populations with the deep history had more time to evolve the ability to perform functions that were useful only in the old environment, and so a higher proportion of their genomes likely

encoded functions that were no longer useful. As a consequence, we reasoned that the lineages with a deep history would be more likely to experience trade-offs when they evolved new functions in the Orthogonal environment, because they could disrupt the ability to perform the functions that were beneficial in the old environment without hindering performance in the new environment. Consistent with our hypothesis, previous research in Avida has shown that antagonistic pleiotropy plays an important role in the adaptive decay of unrewarded functions [69].

Results

Phase I: Evolving founders with shallow and deep history in the old environment.

At the outset of Phase I, every population shared the low fitness of the default organism (fitness = 0.249), which could reproduce but did not perform any computational functions. The 20 different founders for the Phase II populations, which were isolated from the 10 replicate Phase I populations after either shallow or deep history, varied in the fitness they achieved in the Phase I environment. All of them had much higher fitness than the common ancestor (minimum fitness of the 20 founders = 3.826). Not surprisingly, the average fitness measured in the Phase I environment was much greater for the set of founders with deep history than for the set of founders with only a shallow history in that old environment (Figure 3.4). The founders with deep history were also much more variable in fitness than were those with the shallow history (Figure 3.4).

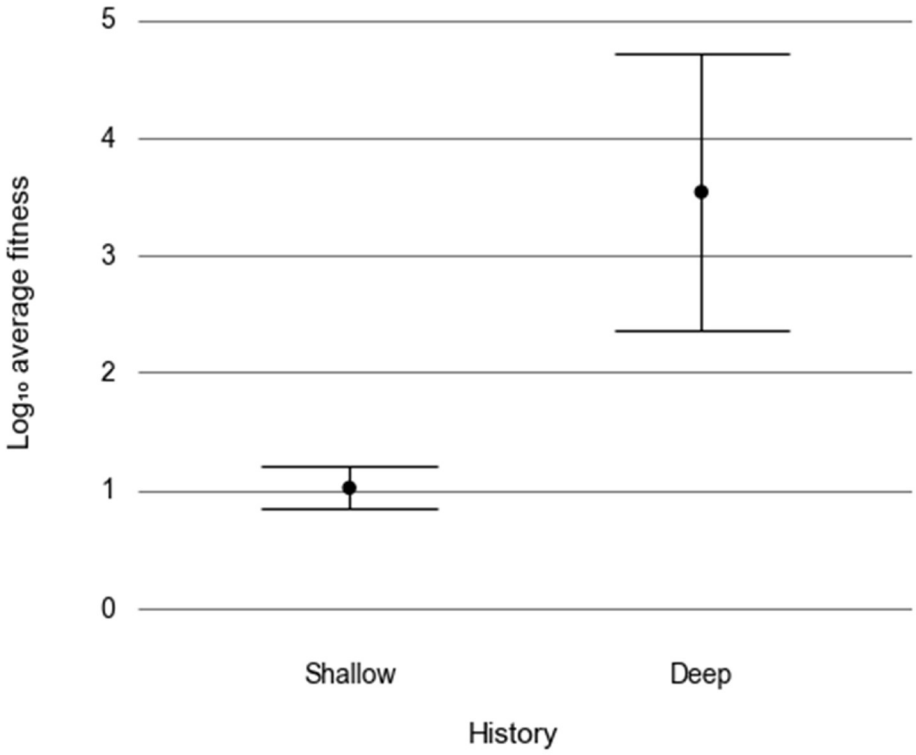


Figure 3.4. Average fitness in the Phase I environment of the Phase II founders with shallow or deep history in that environment
The Phase I environment rewarded organisms for performing 38 Boolean logic functions. The founders used for the Phase II populations were isolated after either a shallow or deep history in the Phase I environment (20,000 or 500,000 updates, respectively). Error bars show 95% confidence intervals. Note the log-transformed scale.

The 20 founders also varied in the number of Boolean logic functions that they could perform, which ranged from 5 to 50. The highest number, in fact, exceeds the number that were rewarded during Phase I, which implies that some functions that were not rewarded occasionally evolved as correlated responses during that first phase. These previously unrewarded functions could be beneficial in a new environment, illustrating the phenomenon that is sometimes called preadaptation. More generally, Figure 3.5 shows that founders with the deep Phase I history evolved to perform significantly more functions than those with a shallow history, and those with the deep history were also more variable in the number of functions they performed.

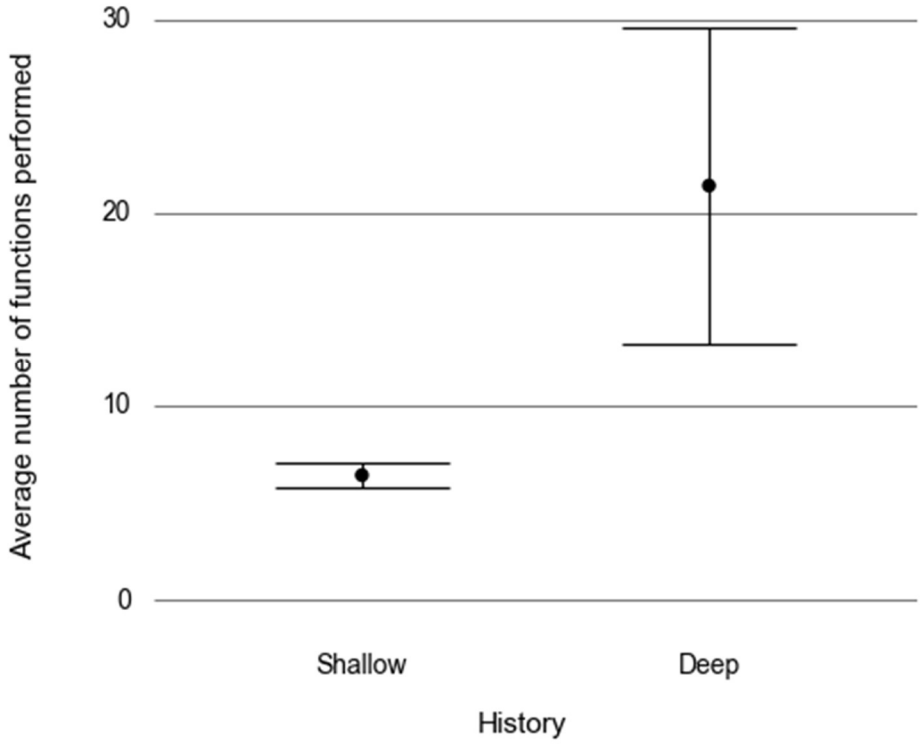


Figure 3.5. Average number of functions performed by the Phase II founders with shallow or deep history
Error bars show 95% confidence intervals.

Phase II: Habitat novelty shapes different responses between founders with shallow and deep history.

In Phase II, we evolved 10 replicate populations from each of the founders from Phase I in two new environments. The *Overlapping* environment was more like the old Phase I environment than the *Orthogonal* environment. In particular, the Overlapping environment rewarded the performance of all 38 functions that the old environment rewarded as well as a suite of 38 new functions. The Orthogonal environment, by contrast, only rewarded the performance of the 38 new functions.

Figure 3.6 shows the average fitness of the founders with shallow and deep history in both the Overlapping and Orthogonal environments. These values thus show the effects only of the prior

evolution that occurred during Phase I. The founders with both shallow and deep history had significantly higher and more variable fitness in the Overlapping environment than in the Orthogonal environment. Also, the founders with deep history had significantly higher and more variable fitness than the founders with shallow history in both environments.

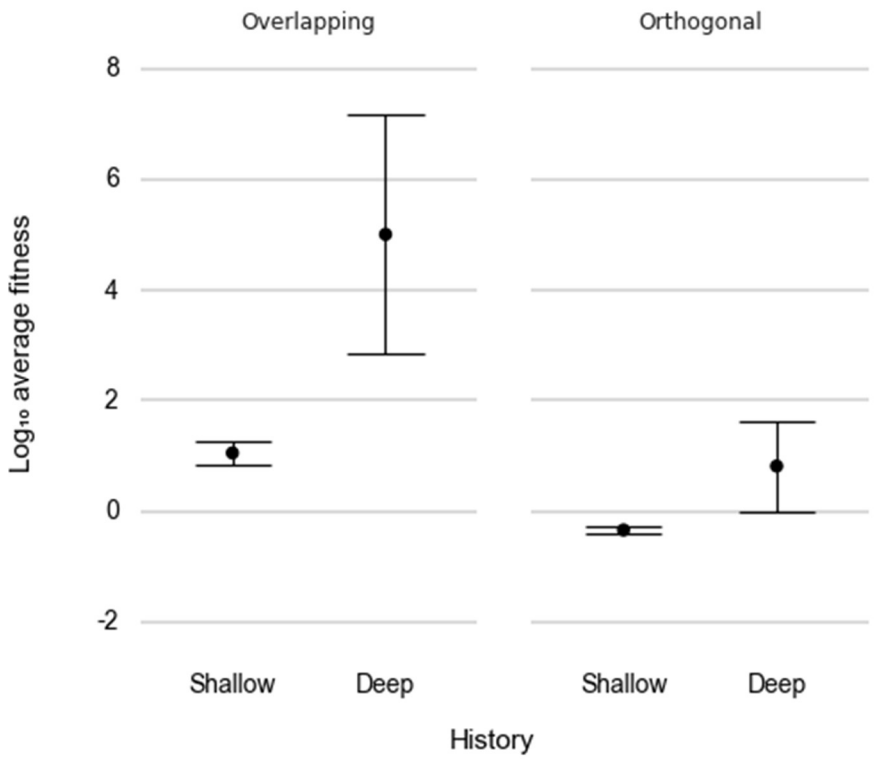


Figure 3.6. Average fitness of Phase II founders with shallow or deep history measured in the new Overlapping and Orthogonal environments
The Overlapping environment rewards organisms for performing the same 38 functions as the old environment as well as an additional 38 other functions. The Orthogonal environment rewards only the other functions. Error bars show 95% confidence intervals. Note the log-transformed scale.

Fitness of founders, gain-of-function genotypes, and final dominant genotypes in the old and new environments.

At the end of Phase II, we found the numerically dominant genotype in each of the 400 populations (i.e.,100 populations for each combination of shallow or deep history and Overlapping or Orthogonal environment). We then identified the gain-of-function genotypes in each lineage

leading from the founder of a Phase II population to the final dominant genotype in that population. A gain-of-function genotype is one that performs at least one logic function that its immediate parent did not. We analyze those genotypes in a later section.

We then plotted the fitness of the founder, each intermediate gain-of-function genotype, and the final dominant genotype from each population in both its ancestral and new environments to examine the cumulative effects of pleiotropy between the two habitats. We generated four figures in this fashion, one for each combination of shallow or deep history, and Overlapping or Orthogonal new environment (Figure 3.7 and Figures 3.13 – 3.15). Each figure has 100 panels, with the 10 rows corresponding to the different founders and the 10 replicate populations for each founder in each row. We present one figure here and provide the others in the Supporting Information. Figure 3.7 shows the fitness of genotypes from lineages that evolved in the Overlapping environment following a shallow history in the ancestral environment. The relatively primitive founders of these populations performed fewer functions at the outset of Phase II than those with a deep history in the ancestral environment (Figure 3.5). Moreover, the new environment rewarded many of the same functions as were rewarded during Phase I. As a consequence, many mutations led to higher fitness in both the old and new environments, as indicated by trajectories moving simultaneously rightward and upward in Figure 3.7. In fact, most lineages evolved overall (cumulative) increases in fitness in both the ancestral and new environments during Phase II, indicating a predominance of synergistic pleiotropy (trade-ups) between the ancestral and Overlapping environments for the populations with shallow history.

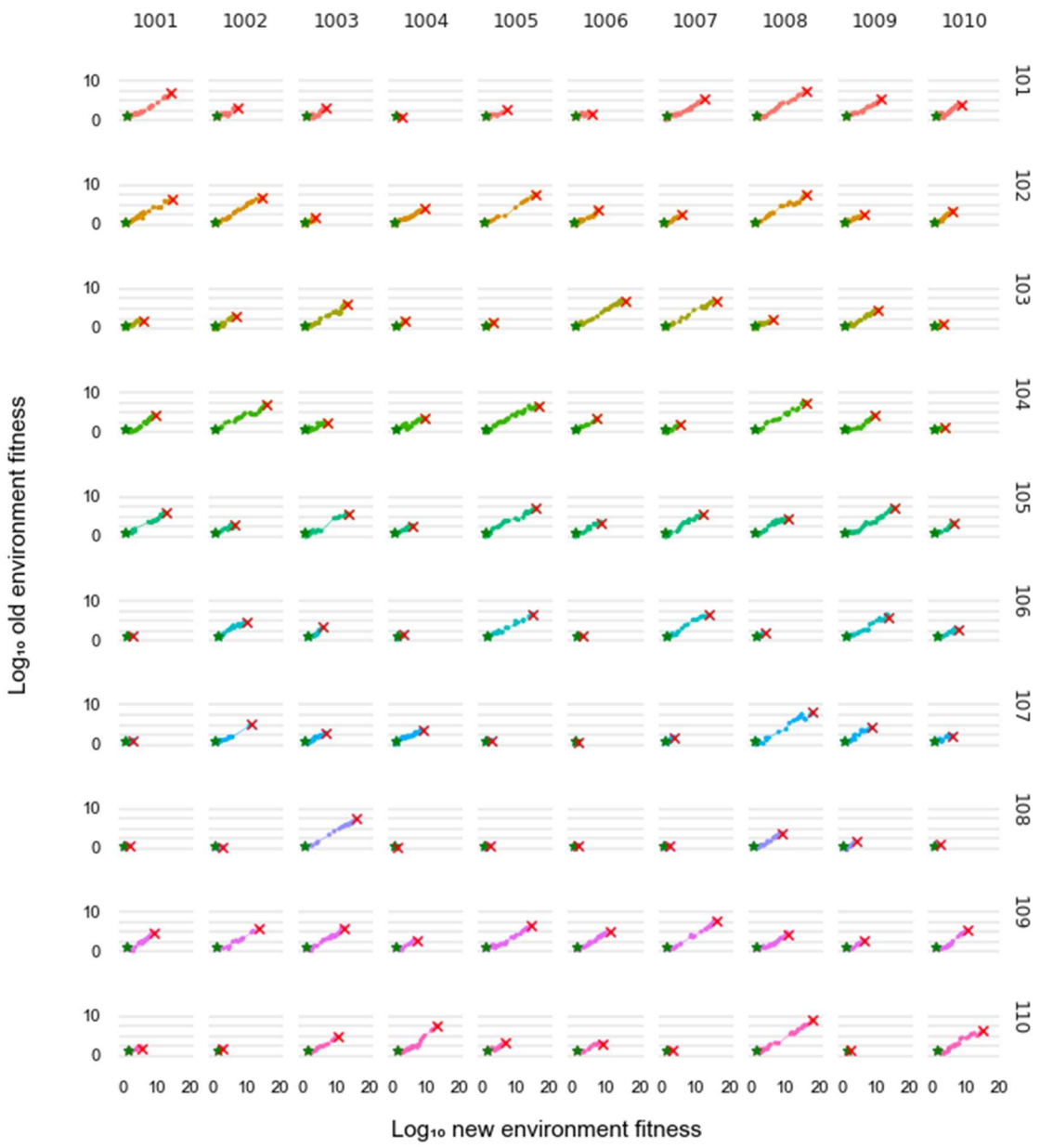


Figure 3.7. Fitness during Phase II for the founders, intermediate gain-of-function genotypes, and final dominant genotypes that evolved in the Overlapping environment with shallow history

This plot shows the fitness of the gain-of-function genotypes (colored dots) in the new (x-axis) and old (y-axis) environments along the line of descent from the founding genotype (green star) to the final dominant genotype (red *x*) for each of the 100 Phase II populations in this treatment. Each row (labels 101-110 at right) uses a different color to show the 10 replicate populations that evolved from the same founder. We used the same 10 numerical seeds (labels 1001-1010 at top) to start the replicate populations derived from each founder. These lineages had shallow history, as their founders evolved for only 20,000 updates in the ancestral environment. Note the log-transformed scales for both axes.

In a similar fashion, Figure 3.13 shows the fitness trajectories for the lines of descent from the 100 populations started from the founders with deep history that then evolved in the Overlapping environment. The founders of these lineages had undergone more adaptive evolution during Phase I than the more primitive founders (Figure 3.6). These deep-history lineages thus evolved fewer new functions during Phase II, and the fitness changes in both the old and new environments tended to be smaller. Moreover, many of these lineages with a deep history actually declined in fitness in the ancestral environment while they increased in fitness in the new environment. Thus, antagonistic pleiotropy (leading to trade-offs) was more common for lineages with a deep history in this environment.

Figure 3.14 shows the fitness trajectories for the 100 lineages started by the founders with shallow history that evolved in the Orthogonal environment during Phase II. Despite the lack of shared resources between the ancestral and new environments, some of these lineages evolved higher fitness in both habitats. However, most lineages showed antagonistic pleiotropy, with sharp fitness declines occurring in the old environment while fitness increased in the new environment (indicated by trajectories moving downward on the y-axis as they move rightward on the x-axis). Figure 3.15 shows the same information for the lineages with deep history evolving in the Orthogonal environment. Most of these lineages also declined in fitness in the ancestral environment as fitness increased in the new environment.

We also note that the difference in trajectories between the lineages with shallow and deep history appears qualitatively to be more distinct for the Overlapping environment (Figure 3.7 and Figure 3.13) than for the Orthogonal environment (Figure 3.14 and Figure 3.15).

Fitness of the final dominant genotype and its founder in the old and new environments.

Figure 3.8 shows the fitness of each final dominant genotype to the fitness of its founder in both the old and new environments. This composite view makes it easier to compare the four treatments in terms of the initial variation among founders as well as the overall direction and magnitude of fitness changes in the old and new environments. Figure 3.9 shows the same information, except we standardized the fitness of each final dominant genotype using the fitness of its founder. That is, we divided each final genotype's fitness by the founder's fitness in each environment, and we then log-transformed that ratio such that a value of *zero* shows no change (fitness ratio = 1). This standardization eliminates the variation among the founders, but it makes it easier to assess the overall differences in the form of the correlated fitness trajectories for each combination of shallow or deep history in Phase I and Overlapping or Orthogonal new environment in Phase II.

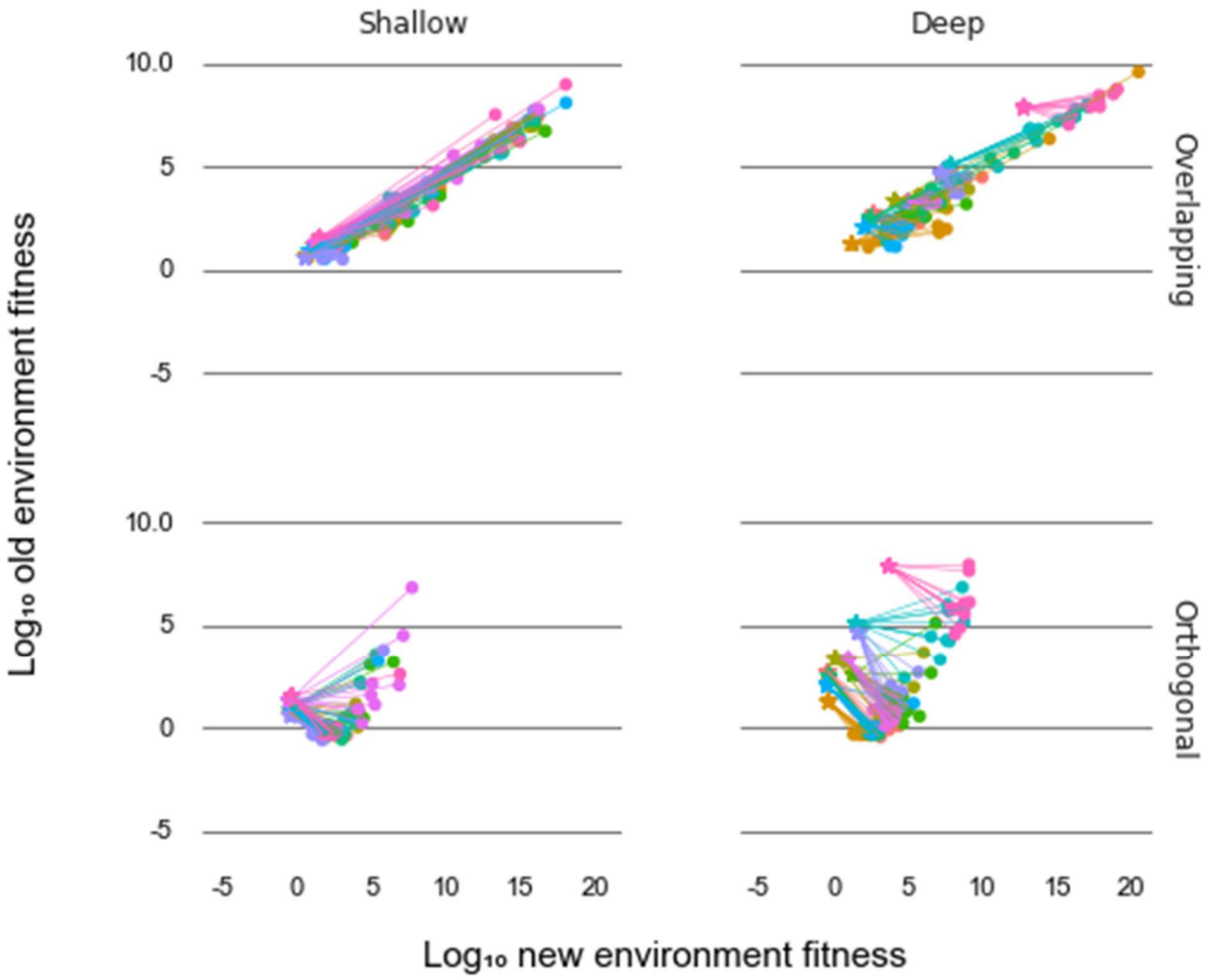


Figure 3.8. Fitness of each population's founder and final dominant genotype in the old and new environments

This plot shows the log-transformed fitness of the final dominant genotype (dots) and the founding genotype (stars) in the old (y-axis) and new (x-axis) environments for all 400 lineages with shallow (left) and deep (right) history during Phase I that evolved in the Overlapping (top) and Orthogonal (bottom) environments during Phase II. The different colors show sets of replicate populations derived from the same founder.

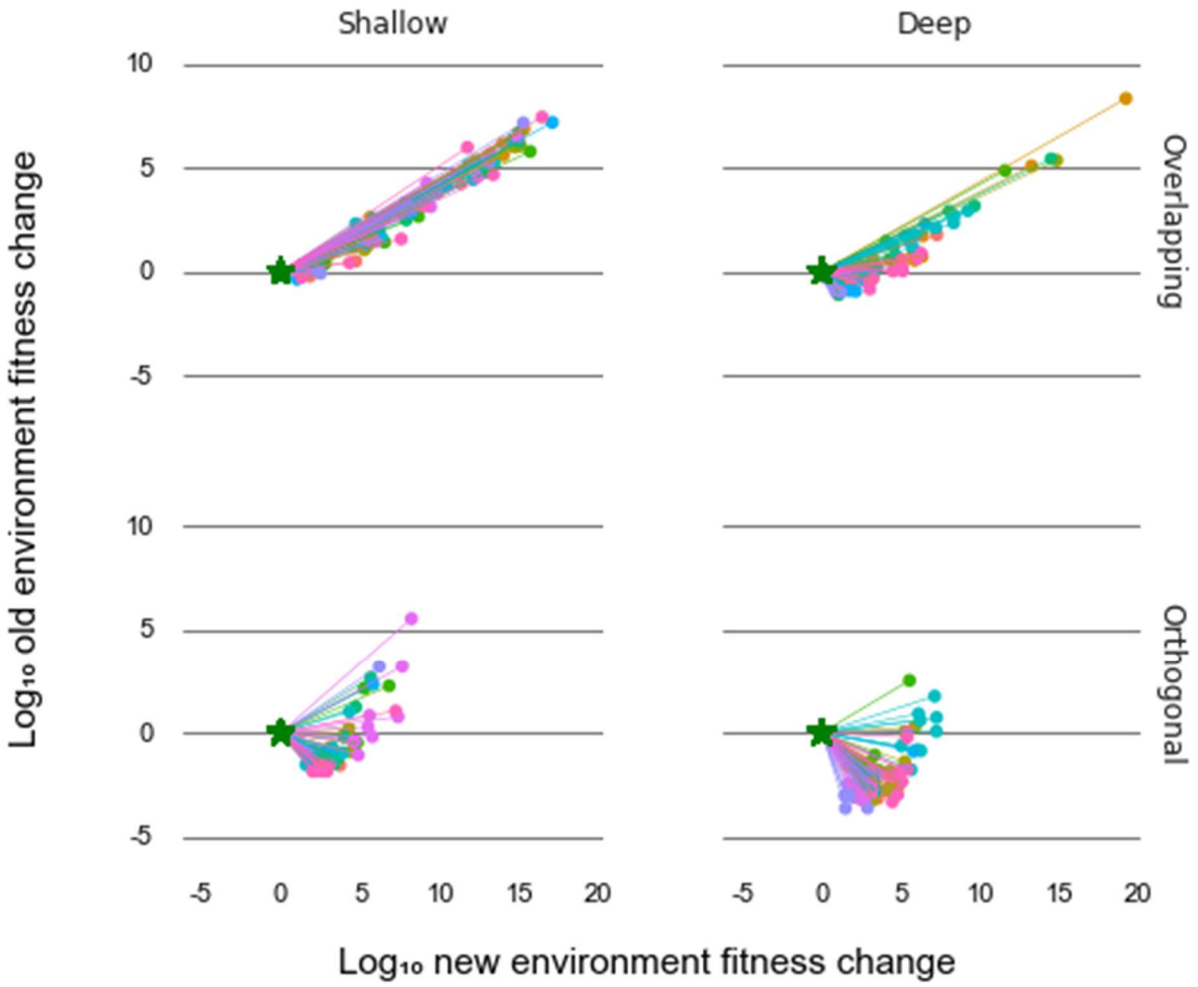


Figure 3.9. Fitness of each final dominant genotype relative to its founder in the old and new environments
This plot shows the log-transformed relative fitness of the final dominant genotype (dots) standardized by the fitness of their founding genotype (star) in the old (y-axis) and new (x-axis) environments for all 400 lineages with shallow (left) and deep (right) history during Phase I that evolved in the Overlapping (top) and Orthogonal (bottom) environments during Phase II. The different colors show sets of replicate populations derived from the same founder.

To summarize these trajectories statistically, we assessed whether the cumulative fitness difference between the final dominant genotype and its founder in the old environment indicated net antagonistic or synergistic pleiotropy. (In all 400 lineages, the final dominant genotype was more fit than its founder in the Phase II environment where it evolved, allowing us to assess the net pleiotropy based on the correlated fitness response in the ancestral environment.) Table 1 shows

the proportion of the final genotypes that showed a trade-off (i.e., reduced fitness in the ancestral environment) indicative of antagonistic pleiotropy for the lineages with shallow and deep history during Phase I that evolved in the Overlapping and Orthogonal environments during Phase II. Table 1 also reports the statistical significance of the effect of the environment for each depth of history, and the effect of the depth of history for each environment.

Environment	History	Lineages	Antagonistic proportion
Overlapping	Shallow	100	0.06
	Deep	100	0.52
Orthogonal	Shallow	100	0.86
	Deep	100	0.91

Environment	Difference	p-value
Overlapping	Shallow vs. Deep	< 0.0001
Orthogonal	Shallow vs. Deep	0.3757

History	Difference	p-value
Shallow	Overlapping vs. Orthogonal	< 0.0001
Deep	Overlapping vs. Orthogonal	< 0.0001

Table 3.1. The proportion of lineages in the four treatments that exhibited cumulative fitness losses in the ancestral environment, indicative of antagonistic pleiotropy
We also ran two-tailed Fisher's exact tests to compare the proportion of lineages that showed antagonistic pleiotropy between the shallow and deep history for each new environment (middle), and between the Overlapping and Orthogonal environments for each depth of history (bottom).

There is a large and highly significant difference in the proportion of lineages showing antagonistic pleiotropy between those with shallow and deep history in the Overlapping environment. In this environment, the lineages with shallow history experienced few trade-offs, as most lineages showed fitness gains even in the ancestral environment. The lineages with the deep history, by contrast, had nearly equal proportions of antagonistic and synergistic pleiotropy. In the Orthogonal environment, the overall proportion of antagonistic pleiotropy was much higher than in the Overlapping environment, and the difference was highly significant for the lineages with

both shallow and deep history. However, the lineages in the Orthogonal environment with shallow and deep history had similar proportions of antagonistic pleiotropy to one another.

Correlated fitness changes in the old and new environments.

We also measured the magnitude as well as the direction of the fitness changes between each final dominant genotype and its founder in both the ancestral and new environments, in order to characterize the correlated responses (Figure 3.10). Figure 3.11 shows the Pearson correlation coefficients and associated 95% confidence intervals for the fitness changes in the two environments for each of the four treatments. We found strong to very strong correlations in all cases. The correlations were significantly stronger in the Overlapping than in the Orthogonal environment at both depths of history. The correlations were also stronger in the case of shallow history than deep history in both environments

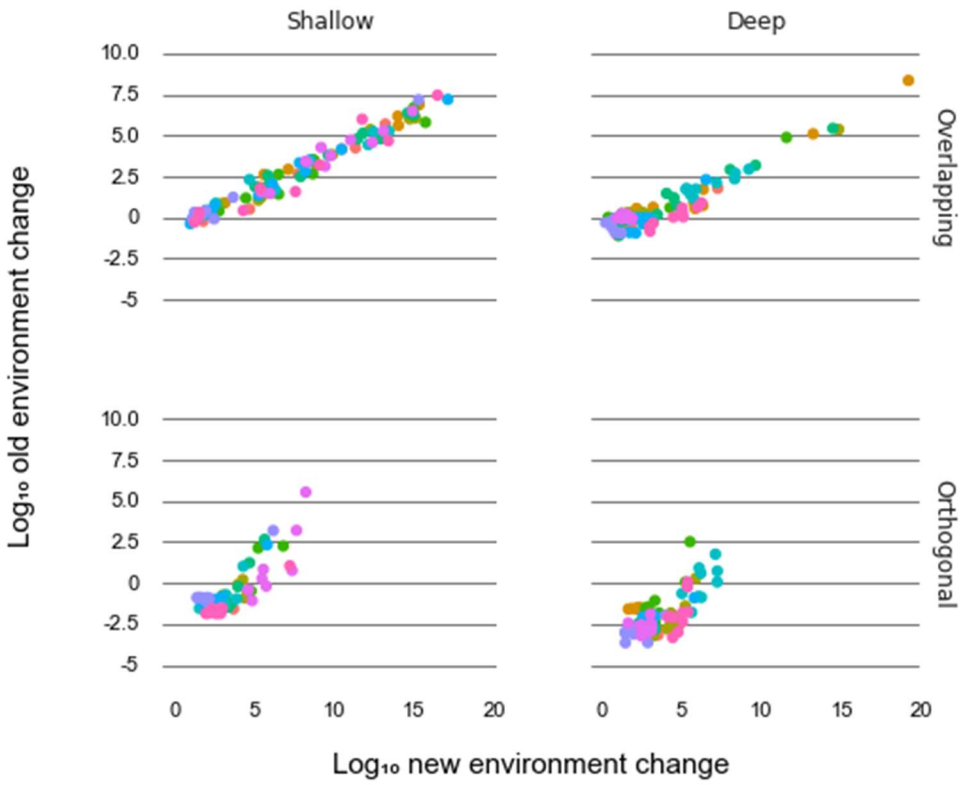


Figure 3.10. Fitness differentials in the old and new environments between each final dominant genotype and its founder

This plot shows the difference in fitness between each final dominant genotype and its founder in the old (y-axis) and new (x-axis) environments from populations with shallow (left) and deep (right) history during Phase I that evolved in the Overlapping (top) and Orthogonal (bottom) environments in Phase II. The different colors show sets of replicate populations derived from the same founder.

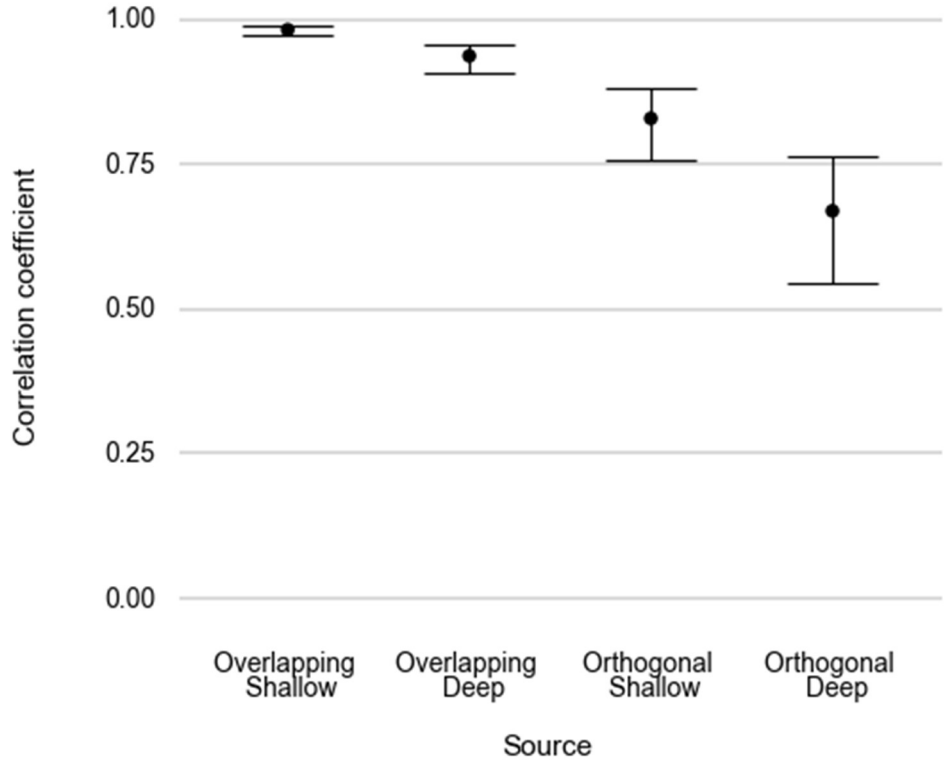


Figure 3.11. Correlation between fitness changes in the old and new environments during Phase II
Each point shows the Pearson correlation coefficient between the fitness changes in the old and new environments when comparing each final dominant genotype to its founder (n = 100 in each case). Error bars show the corresponding 95% confidence intervals.

Contingency and chance both contribute to the among-population variation in pleiotropic effects.

Recall that the 100 populations in each of the four treatments include 10 replicates derived from each of 10 different founders. In this section, we quantify the effects of contingency and chance on the evolution of the direct and correlated fitness responses measured in the new and old environments, respectively. Here, contingency reflects the influence of differences in the founder genotypes on subsequent evolution, whereas chance reflects heterogeneity among independent lineages derived from the same founder. Stated another way, contingency results from the effects of chance events during Phase I on the changes that occurred during Phase II. To that end, we calculated for each lineage a *correlated response metric* as the fitness of the final dominant

genotype relative to its founder in the old environment divided by its fitness relative to its founder in the new environment. This single number encapsulates both the direction and strength of each lineage's fitness change in the old environment as a function of its fitness change in the new environment. Owing to the nested nature of our experiment (10 replicate lineages for each founder from Phase I), we could then perform an analysis of nested variance components to estimate the variance in the correlated response metric within and between groups of lineages that had different founders for each of the four treatments. The effect of the founder group thus reflects historical contingency during Phase II associated with each founder having had an independent *history* during Phase I. The effect of the replicate, by contrast, reflects the influence of *chance* processes (mutation and drift) during Phase II that produce variation in the correlated response metric even among the lineages with the same founder. Figure 3.12 shows the estimated variances associated with the founder (contingency), replicate (chance), and total along with their 95% confidence intervals.

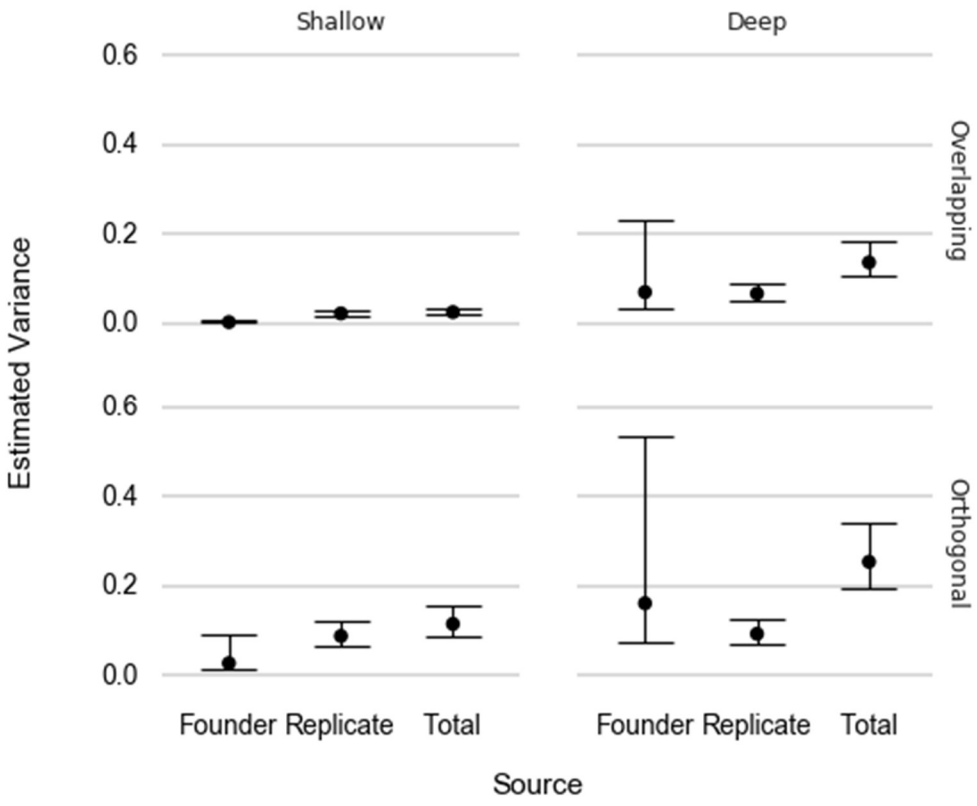


Figure 3.12. History and chance contribute to variation in correlated responses
We estimated the variance in the correlated response metric (see text) for lineages with different proximate ancestors, each with unique history from Phase I (founder), and the variance due to chance between lineages with the same founder genotype (replicate), for each combination of shallow (left) or deep (right) history in Phase I and Overlapping (top) or Orthogonal (bottom) environment during Phase II. Error bars show 95% confidence intervals, all of which exclude 0.

Both history and chance explained significant amounts of variation in the correlated response metric for all four treatments. In both environments, chance (variation among replicates with the same founder) contributed significantly more to this metric than contingency (variation attributable to the different founders) when history was shallow. However, when the Phase I history was deep, contingency's effect was as large as that of chance in both environments. The results were qualitatively similar when we performed this analysis using a different metric, namely the proportion of gain-of-function mutations along each line of descent showed antagonistic pleiotropy between the old and new environments (Figure 3.16). Thus, Figure 3.12 examines both the

direction and magnitudes of the cumulative changes in relative fitness in the old and new environments in each lineage, whereas Figure 3.16 focuses on the proportion of individual gain-of-function mutations in each lineage that exhibit antagonistic pleiotropy (i.e., trade-offs).

These patterns imply that the novelty of the environment and the depth of history together can shape the genetic architecture and the resulting prevalence of antagonistic pleiotropy in these digital organisms. In the Overlapping environment, lineages tend to retain adaptations that arose during Phase I owing to the positive fitness correlation that results from many of the same functions being rewarded in both the old and new environments. Maintaining ancestral adaptations that are still under positive selection contributes to greater variation among groups with different founders, and this effect is magnified as history becomes deeper. By contrast, in the Orthogonal environment, there is no continuity in the specific logic functions that are rewarded in Phases I and II. Moreover, the effects of contingency and chance are relatively large, in comparison to the Overlapping environment, even after a shallow history during Phase I. Thus, the effects of contingency and chance are more similar, at least in relative terms, for shallow and deep history in the Orthogonal environment than in the Overlapping environment.

Discussion

In *Wonderful Life* Gould speculated, "perhaps genetic systems do 'age' in the sense of becoming 'less forgiving of major restructuring'" [52]. Our results suggest that the extent of habitat novelty plays an important role in determining how difficult it is for old genomes to "learn new tricks" (see also [30]). We found constraints associated with having founders with deep history when populations evolved in the Overlapping environment, a new habitat that was similar to the ancestral environment. While there was less antagonistic pleiotropy with the old environment overall in the Overlapping environment than in the more dissimilar Orthogonal environment, there was a much

stronger effect of deep history in the Overlapping environment. In particular, there was a shift in the Overlapping environment from mostly synergistic pleiotropy (trade-ups) after shallow history to a fairly even balance between antagonistic and synergistic pleiotropy after deep history in the ancestral environment (Table 1). The prevalence of synergistic pleiotropy in the case of shallow history occurred because the more primitive founders had evolved fewer functions that would continue to be rewarded during Phase II, leading to more gain-of-function mutations that would be beneficial in both habitats. Thus, in the Overlapping environment, we saw a dramatic shift in the form of pleiotropy (from synergistic to antagonistic) as the depth of history during Phase I was extended.

Intuitively, it might seem that having more time to adapt to the more similar Overlapping environment would provide an advantage for the lineages with deep history. Indeed, these lineages had a head start, with founders that had higher (and more variable) fitness in the new environment. But as a consequence, there were fewer remaining opportunities for the lineages with deep history to make further gains during Phase II in this environment. As a further consequence, historical contingency and chance effects generated more variation in correlated responses (Figure 3.12) and the proportion of gain-of-function mutations with antagonistic pleiotropic effects (Figure 3.16) for those lineages with a deep history in the Overlapping environment. Taken together, these effects help to explain why adaptation was more constrained for populations with a deep history in the Overlapping environment (Figure 3.3).

In the Orthogonal environment that was dissimilar to the ancestral environment, the overall prevalence of antagonistic pleiotropy (tradeoffs) was much higher. However, the difference in antagonistic effects between shallow and deep lineages in this new environment was small and not statistically significant (Table 1). We used the same founders in both Phase II environments, and

so we know that the founders with a deep history in Phase I could perform more functions in the Orthogonal environment as well. However, any instructions used exclusively for functions that were rewarded in the old environment were unconstrained in the Orthogonal environment, because performing those functions no longer conferred a benefit. In the Orthogonal environment, the lineages with shallow and deep histories explored opportunities of more or less equivalent breadth, unlike in the Overlapping environment. In effect, the genetic constraints of "old age" (deep history) postulated by Gould were reduced for the lineages adapting to the more novel conditions.

Future Directions

There are additional issues we could address related to contingency and convergence in our virtual "long-term" evolution experiment, as well as other approaches we could employ. One example involves analyzing instances of homoplasy. Homoplasy refers to the independent evolution of similar traits in multiple lineages, where the traits were absent in the lineages' most recent common ancestor. Powell acknowledged that biologists typically use *parallel* evolution when referring to homoplastic traits in closely related lineages, and they use *convergent* evolution when referring to homoplasy between distantly related lineages. But this distinction is somewhat arbitrary. Instead, Powell suggested restricting the term parallelism to those cases in which the similar (homoplastic) trait directly involves a "developmental homology", which we might broadly interpret to include the same or similar mechanistic basis in terms of the mapping of genotype to phenotype. Our ability, in Avida, to isolate every genotype along a line of descent, and then map the performance of particular functions to specific genomic instructions, might enable us to implement a version of Powell's "screening off" test to distinguish between parallelism and convergence in this fashion [70–72]. This analysis would involve identifying particular functions that evolved during Phase II in multiple lineages derived from the same founder. We could then determine when the ability to

perform these functions arose and, moreover, whether they repeatedly emerged from the same precursor functions, which would show developmental homology in the sense of Powell. Parallel evolution is indicated when a necessary precursor function is inherited from the founder. But the homoplasy is convergent when the necessary precursor functions also evolved independently and led to the development of the new function even if some earlier, commonly-inherited precursors are knocked out ("screened off"). In essence, this test seeks to distinguish between parallelism that results from historically contingent constraints internal to the organism itself, on the one hand, and convergence driven by adaptation to shared environmental factors external to the organism, on the other hand, rather than relying on arbitrary and ill-defined measures of phylogenetic divergence [70–72]. By performing such an analysis in our study system, we might be able to determine whether and how the levels of parallelism and convergence depend on the depth of history and the extent of environmental change.

Xie et al. identified permissive and restrictive mutations that arose during the experimental evolution of proteins [3]. In the context of our study system, a permissive mutation would be one that allows subsequent mutations to produce a new function on a genetic background with deep history, but not on a background with shallow history from the same lineage. A restrictive mutation, on the other hand, would prevent a mutation that yielded a new function on a background with shallow history from generating the same new function on a background with deep history. By analyzing not only the fitness and functionality of genotypes along a line of descent, but also the effects of every one-step mutation on each genotype's properties, we might be able identify founder-specific mutations (i.e., ones that occurred during Phase I) that interact epistatically with gain-of-function mutations that arose during Phase II [3,48]. We could then determine how the

balance of permissive and restrictive mutations depends on the depth of history and the extent of habitat novelty.

In this study, we focused on the evolution of new functions. However, Ostrowski et al. used Avida to analyze how antagonistic pleiotropy and mutation accumulation contributed to loss-of-function mutations (adaptive decay) in specialized descendants of generalist ancestors that had evolved in an environment that rewarded only a single function [69]. Mutation accumulation is a non-adaptive process that occurs when unused functions decay as the result of relaxed selection. It should similarly be possible to evaluate how antagonistic pleiotropy and mutation accumulation contributed to losses of function (rather than gains of function) in our experiments. In our system, mutation accumulation would likely be more frequent in the Orthogonal environment, where functions that evolved during Phase I could be more readily lost without adversely affecting fitness in the new environment.

There are countless ways to configure the virtual world of Avida. However, there are some practical limitations as well, and while it instantiates evolution as a general process it is not meant to "simulate" biological organisms [73]. For instance, populations are much smaller, and per-site mutation rates much higher, when compared to bacteria. On the other hand, Avida shares several important features with the microbial systems often used in evolution experiments including rapid generations, asexual reproduction, and the ability to store and later revive sampled organisms. In any case, the tremendous flexibility of Avida makes it possible to extend our investigation in ways that would not be feasible using a biological model system. Our "wish list" is inspired by previous research in Avida that implements features such as resource limitation [50,74–76], quorum sensing [77–82], predation [83–85], and sexual reproduction [86–89]. Future experiments could examine how these features affect the interaction of history and habitat novelty. These alternate worlds

would likely produce different interactions of adaptation, chance, and history than those we have reported (i.e. Figure 3.3 and [66]). The challenge, then, would be to identify possible generalities and put them to the test.

In this study and its predecessor [66], we have documented several interesting effects of the depth of history's footprint, the extent of environmental change, and their interaction on patterns and dynamics of evolution. For example, deepening history constrained adaptation and radically changed the mix of antagonistic and synergistic pleiotropy in the Overlapping environment, whereas these effects were absent or diminished in the Orthogonal environment. One factor that may have contributed to these differences is that neither depth of history during Phase I was sufficient to allow the populations to reach a quasi-equilibrium state. Future experiments could establish and examine much deeper histories, perhaps even sufficient for populations to reach local fitness peaks. With far more extensive divergence between ancestors than is typically possible in an evolution experiment, the correlation between fitness changes in the old and new environments might be substantially weaker than, or even reversed relative to, what we observed after 500,000 updates (around 65,000 generations), even in the Overlapping environment (Figure 3.11). We can imagine an experiment in Avida with a Phase I duration of a billion updates that could measure evolutionary trajectories over tens of millions of generations. This extremely long-term evolution experiment would be difficult to sustain in vivo, but it should be possible in this digital system with sufficient computational resources. We would, of course, want to examine how the additional "epochs" of divergence influence the subsequent ability of descendant populations to evolve in new habitats (Phase II). But the results of such an extensive Phase I would be interesting in their own right as well. As a point of comparison, the *E. coli* Long-Term Evolution Experiment (LTEE) is the longest running experiment in terms of the number of generations. The

results to date from the LTEE suggest that even in a constant and simple environment the dynamics of long-term adaptation follow a power-law model, in which fitness has no upper limit, although the rate of improvement declines over time [68,90]. It would be interesting to see how evolution plays out over the very long run in our system, in part because of certain difficulties with making and testing predictions based on the shorter periods investigated so far [91].

Materials and Methods

Avida system and original experiment

We performed a two-phase replay experiment using the program Avida (version 2.14). Avida is a platform for performing evolution experiments with digital organisms. Avida is available for free [95]. Each organism has a genome that encodes its behaviors including the ability to replicate and perform Boolean logic operations. A genome in Avida is circular, with 26 instructions possible at each location. The default ancestral organism has a genome with 100 instructions. It can reproduce but not perform any other functions. An organism reproduces by copying its genome into a new memory location. Mutations occur at random during replication and post-replicative division. We set the point-mutation rate during replication at 0.007 per instruction and the corresponding rate of indels (insertions or deletions of single instructions) at 0.00005 per instruction. Indels also occurred at a rate of 0.1 per post-replicative division. Our configuration assigned each new organism to a random location (in essence, a well-mixed population), where it replaced the current occupant (which therefore died), resulting in genetic drift. An organism could also die of old age after executing a total number of instructions that exceeded 20 times its genome length; such deaths generally only occur in organisms with highly deleterious mutations that generate endless or near-endless execution loops. We set the maximum population size at 3,600 organisms in both phases.

The experimenter defines environments in Avida by configuring the rewards for various Boolean logic functions (sometimes also called tasks). Avida rewards an organism for performing a function if it outputs a 32-bit string that matches the expected result of performing the associated bit-wise logic function on one, two, or three input strings, provided that the environment is configured to reward that function. The organism that executed the function then receives extra CPU cycles, which allow it to execute additional instructions that may help it reproduce faster than

its competitors. We configured Avida to only reward the performance of any given function once per organism. We defined our environments using subsets of the 77 two- and three-input functions available in Avida. We excluded the most difficult two-input function (equals, or EQU). We divided the remaining 76 functions into three sets corresponding to even-numbered functions, odd-numbered functions, and all functions.

In Phase I, we evolved 10 replicate populations from the default ancestral organism for 500,000 updates (the unit of time in Avida). The environment used for Phase I rewarded the 38 even-numbered functions. During Phase II, we evolved 10 replicate populations from the dominant organism isolated from each Phase I population, and we did so in each of two new environments. The Overlapping environment rewarded all 76 functions, including 38 in common with the Phase I environment. The Orthogonal environment rewarded only the odd-numbered functions, and therefore it shared no resources with the old environment. We isolated the dominant (i.e., most abundant) genotype from each Phase I population at two times, which correspond to shallow (20,000 updates) and deep (500,000 updates) history. These 20 genotypes then served as founders for the 4000 Phase II populations [66].

Analysis

We began this study by re-running Phase II of the experiment. This generated new data with the same configuration files as used in our previous work [66]. We did so because the earlier runs used an older version of Avida that did not allow some of the analyses to be performed that were needed for the present study. These re-runs yielded estimates of the contributions of adaptation, chance, and history to fitness in all four treatments (two new environments, and two depths of history) that are nearly indistinguishable (Figure 3.3) from our previous results [66]. Two of the re-runs produced errors, and we replaced them by using new starting seeds.

To assess the initial variation among the Phase II populations, we analyzed the effect of having proximate ancestors (founders) with a shallow or deep history during Phase I. We measured the fitness of all 20 founders (10 Phase I lineages x 2 depths of history) in the Phase I environment, and we counted the total number of functions performed by each founder. We estimated the average fitness (Figure 3.4) and average total number of functions performed (Figure 3.5) for both sets of founders (i.e., with shallow and deep history). We also measured the average fitness of the two sets of founders in both of the new environments used for Phase II (Figure 3.6).

Our re-runs included instructions for Avida to run in "analyze mode" following the completion of Phase II. In this mode, Avida loads specified genotypes into specified environments and then runs tests that generate additional data not provided in the normal operating mode. In particular, we configured Avida to isolate and test each genotype on the line of descent leading from the founder to the final dominant genotype for each Phase II population. These tests provided the fitness of each genotype in both the old (Phase I) and new (Phase II) environments along with the functions it could perform. To identify gain-of-function mutations, we discarded any genotype that did not perform at least one function that its parent did not.

We then plotted the fitness of the founder, each gain-of-function genotype, and the final dominant genotype from each lineage in the old and new environments for each of the four treatments to visualize the effects of pleiotropy (Figure 3.7, Figure 3.13, Figure 3.14, and Figure 3.15). In Avida, a genotype's fitness is its rate of "energy" acquisition divided by the energetic cost of producing an offspring, where energy is measured in terms of the execution of instructions. The rate of execution is increased above a baseline rate when an Avidian performs logic functions that are rewarded in its current environment, while the cost of producing an offspring is the number of instructions that the organism executes during its lifetime leading up to a successful reproduction.

This number is much greater than the organism's genome length (the number of instructions in its genome) because many instructions are executed repeatedly while performing functions and replicating the genome.

We compared the fitness of the final dominant genotype in each lineage to its founder in both the old and new environments (Figures 3.8 and 3.9). We calculated the proportion of lineages that showed net antagonistic pleiotropy between fitness changes in the two environments (Table 1 top). We performed Fisher's exact tests to compare the proportion of lineages showing net antagonistic pleiotropy between the shallow and deep history for each new environment (Table 1 middle), and between the Overlapping and Orthogonal environments for each depth of history (Table 1 bottom).

We also examined the quantitative relationship between fitness changes in the old and new environments for each lineage (Figure 3.10). For each of the four treatments, we calculated the Pearson correlation coefficient between the relative fitness changes in the two environments along with the 95% confidence intervals (Figure 3.11).

We calculated for each lineage a correlated response metric, which we defined as the fitness of the final dominant genotype relative to its founder in the old environment divided by its fitness relative to its founder in the new environment. This metric thus encapsulates both the direction and strength of each lineage's fitness change in the old environment as a function of its fitness change in the new environment. Given our nested experimental design, we then estimated for each treatment the variance components that reflect the effects of contingency (i.e., different founders) and chance (i.e., heterogeneity among replicate populations derived from the same founder) for this metric (Figure 3.12). We computed approximate 95% confidence intervals for these variance components based on the chi-square distribution [96]. Finally, we performed a similar analysis for

the proportion of gain-of-function mutations along each line of descent that showed antagonistic pleiotropy for fitness between the old and new environments (Figure 3.16).

We used Python (version 3) for all of these analyses. All of our configuration files, data, and analysis scripts are available online [97].

APPENDIX

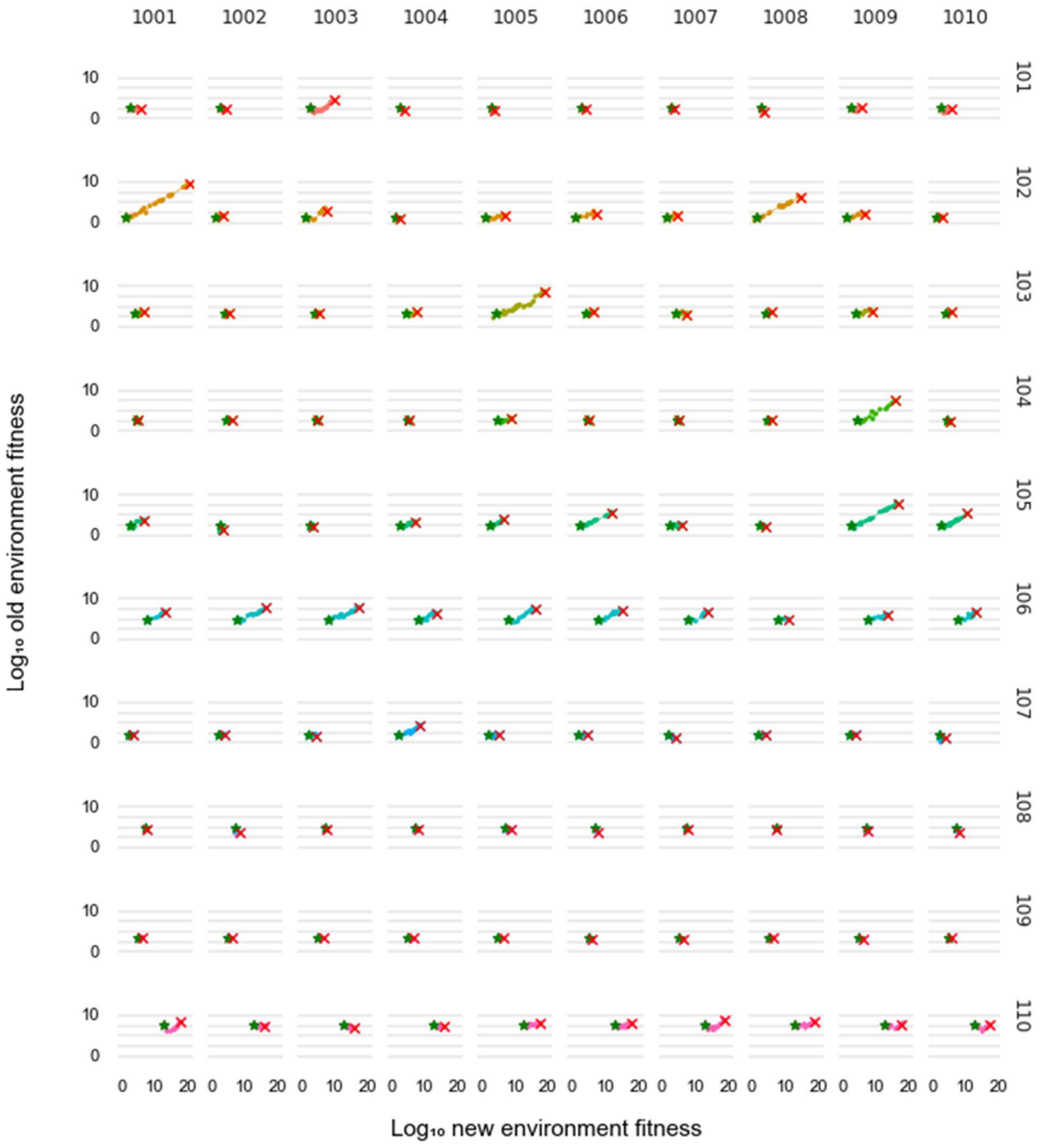


Figure 3.13. Fitness during Phase II for the founders, intermediate gain-of-function genotypes, and final dominant genotypes that evolved in the Overlapping environment with deep history

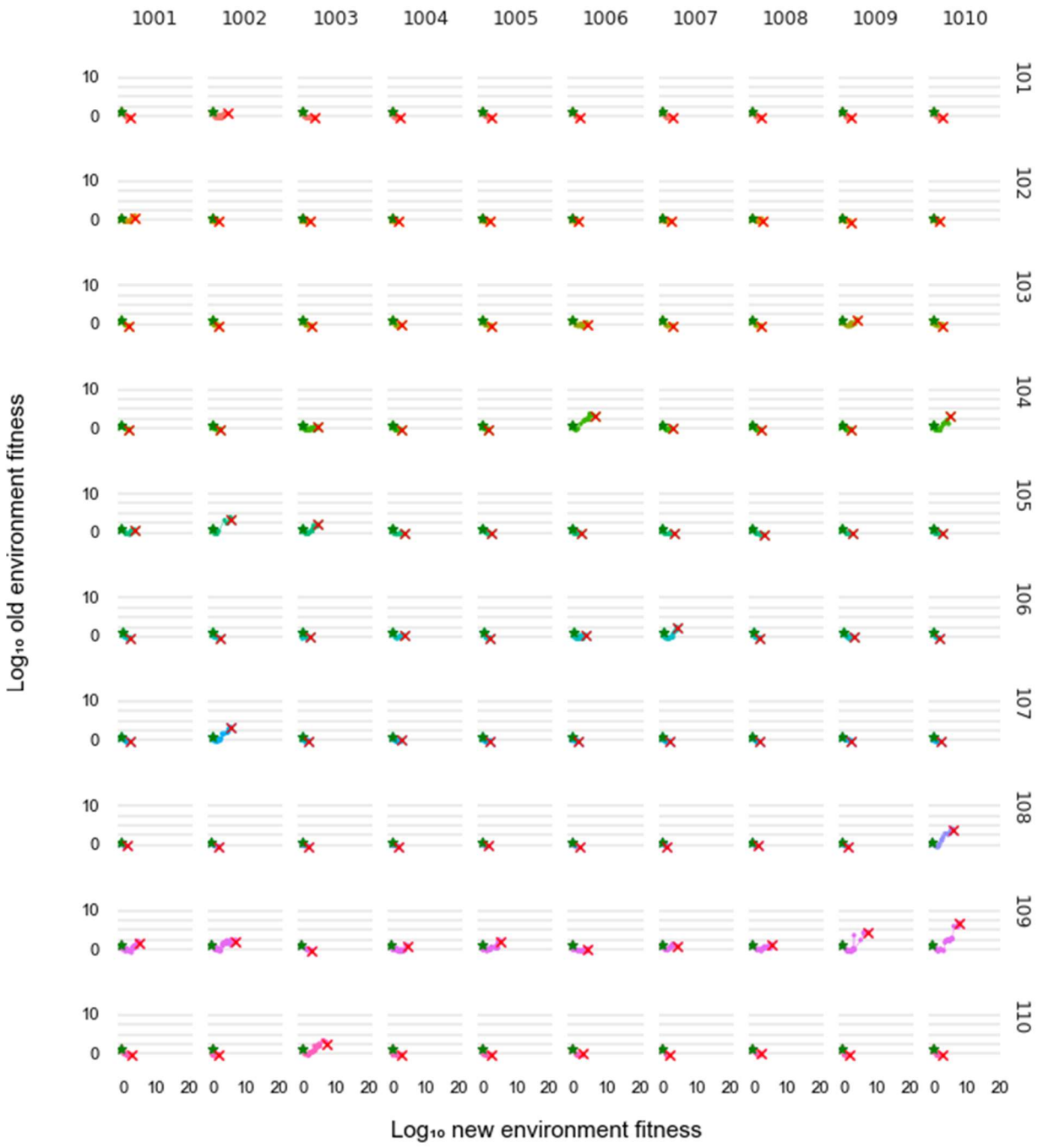


Figure 3.14. Fitness during Phase II for the founders, gain-of-function intermediate genotypes, and final dominant genotypes that evolved in the Orthogonal environment with shallow history

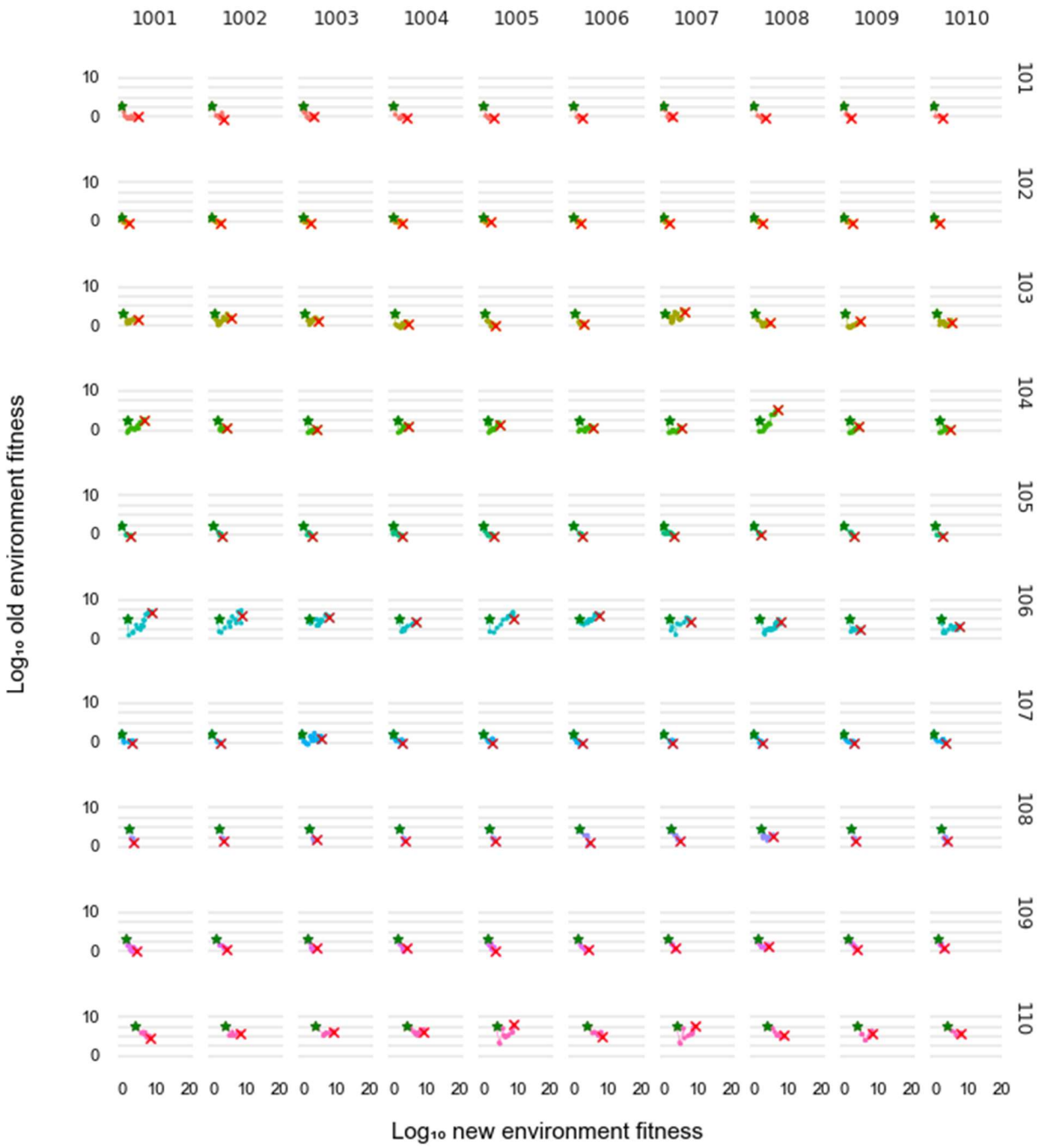


Figure 3.15. Fitness during Phase II for the founders, intermediate gain-of-function genotypes, and final dominant genotypes that evolved in the Orthogonal environment with deep history

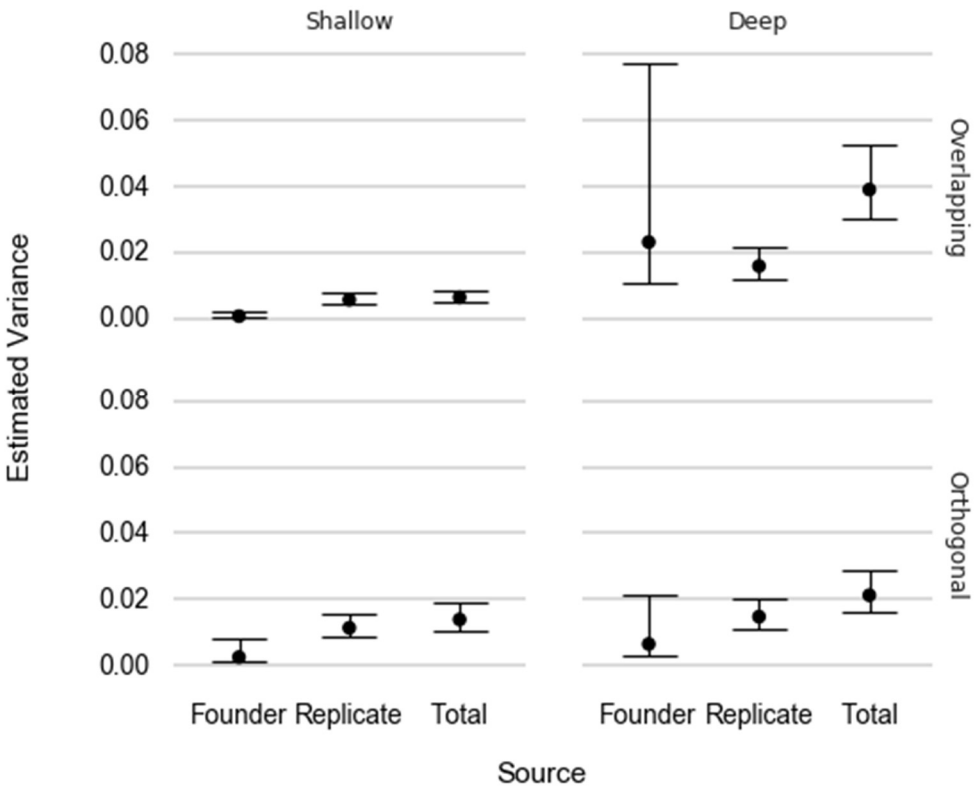


Figure 3.16. History and chance contribute to variation in the proportion of antagonistic gain-of-function mutations

We estimated the variance in the proportion of gain-of-function mutations that showed antagonistic pleiotropy for fitness in the old and new environments for lineages with different proximate ancestors, each with unique history from Phase I (founder), and the variance due to chance between lineages with the same founder genotype (replicate), for each treatment. Error bars show 95% confidence intervals, all of which exclude 0.

REFERENCES

REFERENCES

1. Blount ZD, Lenski RE, Losos JB. Contingency and determinism in evolution: Replaying life's tape. Science. 2018;362(6415):1–10.

2. Sane M, Miranda JJ, Agashe D. Antagonistic pleiotropy for carbon use is rare in new mutations. Evolution. 2018;72(10):2202–2213.

3. Xie VC, Pu J, Metzger BP, Thornton JW, Dickinson BC. Contingency and chance erase necessity in the experimental evolution of ancestral proteins. eLife. 2021;10:e67336.

4. Futuyma DJ, Moreno G. The evolution of ecological specialization. Annu Rev Ecol Syst. 1988;19(1):207–233.

5. Rees M. Trade-offs among dispersal strategies in British plants. Nature. 1993;366(6451):150–152.

6. Agrawal AA, Conner JK, Rasmann S. Tradeoffs and negative correlations in evolutionary ecology. Evolution since Darwin. 2010;150:243–268.

7. Ferenci T. Trade-off mechanisms shaping the diversity of bacteria. Trends Microbiol. 2016;24(3):209–223.

8. Tilman D. Causes, consequences and ethics of biodiversity. Nature. 2000;405(6783):208–211.

9. Bohannan BJ, Kerr B, Jessup CM, Hughes JB, Sandvik G. Trade-offs and coexistence in microbial microcosms. Antonie Van Leeuwenhoek. 2002;81(1):107–115.

10. Litchman E, Edwards KF, Klausmeier CA. Microbial resource utilization traits and trade-offs: Implications for community structure, functioning, and biogeochemical impacts at present and in the future. Front Microbiol. 2015;6:254.

11. Kirkwood TBL. Understanding the odd science of aging. Cell. 2005;120(4):437–447.

12. Stearns SC. The evolution of life history traits: A critique of the theory and a review of the data. Annu Rev Ecol Syst. 1977;8:145–171.

13. Stearns SC. Trade-offs in life-history evolution. Funct Ecol. 1989;3(3):259–268.

14. Stearns SC. Life history evolution: Successes, limitations, and prospects. Naturwissenschaften. 2000;87(11):476–486.

15. Zera AJ, Harshman LG. The physiology of life history trade-offs in animals. Annu Rev Ecol Syst. 2001;32(1):95–126.

16. Sgrò CM, Hoffmann AA. Genetic correlations, tradeoffs and environmental variation. Heredity. 2004;93(3):241–248.

17. Dessau M, Goldhill D, McBride RL, Turner PE, Modis Y. Selective pressure causes an RNA virus to trade reproductive fitness for increased structural and thermal stability of a viral enzyme. PLoS Genet. 2012;8(11).

18. Goldhill DH, Turner PE. The evolution of life history trade-offs in viruses. Curr Opin Virol. 2014;8:79–84.

19. Burmeister AR, Sullivan RM, Lenski RE. Fitness costs and benefits of resistance to phage lambda in experimentally evolved *Escherichia coli*. In: Evolution in Action: Past, Present and Future. Springer; 2020.

20. Elena SF, Lenski RE. Evolution experiments with microorganisms: The dynamics and genetic bases of adaptation. Nat Rev Genet. 2003;4(6):457–469.

21. Bono LM, Smith Jr LB, Pfennig DW, Burch CL. The emergence of performance trade-offs during local adaptation: Insights from experimental evolution. Mol Ecol. 2017;26(7):1720–1733.

22. Burmeister AR, Turner PE. Trading-off and trading-up in the world of bacteria–phage evolution. Curr Biol. 2020;30(19):R1120–1124.

23. Kneitel JM, Chase JM. Trade-offs in community ecology: Linking spatial scales and species coexistence. Ecol Lett. 2004;7(1):69–80.

24. Clark AG. Sperm competition and the maintenance of polymorphism. Heredity. 2002;88(2):148–153.

25. Ostrowski EA, Rozen DE, Lenski RE. Pleiotropic effects of beneficial mutations in *Escherichia coli*. Evolution. 2005;59(11):2343–2352.

26. Duffy S, Turner PE, Burch CL. Pleiotropic costs of niche expansion in the RNA bacteriophage phi 6. Genetics. 2006;172(2):751–757.

27. Reusch TBH, Wood TE. Molecular ecology of global change. Mol Ecol. 2007;16(19):3973–3992.

28. Stearns FW. One hundred years of pleiotropy: A retrospective. Genetics. 2010;186(3):767–773.

29. Østman B, Hintze A, Adami C. Impact of epistasis and pleiotropy on evolutionary adaptation. Proc R Soc Lond B Biol Sci. 2012;279(1727):247–256.

30. McGee LW, Sackman AM, Morrison AJ, Pierce J, Anisman J, Rokyta DR. Synergistic pleiotropy overrides the costs of complexity in viral adaptation. Genetics. 2016;202(1):285–295.

31. Frachon L, Libourel C, Villoutreix R, Carrère S, Glorieux C, Huard-Chauveau C, et al. Intermediate degrees of synergistic pleiotropy drive adaptive evolution in ecological time. Nat Ecol Evol. 2017;1(10):1551–1561.

32. Schreiber SJ, Patel S, terHorst C. Evolution as a coexistence mechanism: Does genetic architecture matter? Am Nat. 2018;191(3):407–420.

33. Burmeister AR, Fortier A, Roush C, Lessing AJ, Bender RG, Barahman R, et al. Pleiotropy complicates a trade-off between phage resistance and antibiotic resistance. Proc Natl Acad Sci U S A. 2020;117(21):11207–11216.

34. Dutta A, Hartmann FE, Francisco CS, McDonald BA, Croll D. Mapping the adaptive landscape of a major agricultural pathogen reveals evolutionary constraints across heterogeneous environments. ISME J. 2021;15(5):1402–1419.

35. Tienderen PHV. Evolution of generalists and specialists in spatially heterogeneous environments. Evolution. 1991;45(6):1317–1331.

36. Wilson DS, Yoshimura J. On the coexistence of specialists and generalists. Am Nat. 1994;144(4):692–707.

37. Gilchrist GW. Specialists and generalists in changing environments. I: Fitness landscapes of thermal sensitivity. Am Nat. 1995;146(2):252–270.

38. Bernays EA, Funk DJ. Specialists make faster decisions than generalists: Experiments with aphids. Proc R Soc Lond B Biol Sci. 1999;266(1415):151-156.

39. Kassen R. The experimental evolution of specialists, generalists, and the maintenance of diversity. J Evol Biol. 2002;15(2):173–190.

40. Julliard R, Clavel J, Devictor V, Jiguet F, Couvet D. Spatial segregation of specialists and generalists in bird communities. Ecol Lett. 2006;9(11):1237–1244.

41. Luria SE, Delbrück M. Mutations of bacteria from virus sensitivity to virus resistance. Genetics. 1943;28(6):491–511.

42. Jacob F. Evolution and tinkering. Science. 1977;196(4295):1161–1166.

43. Goldsmith TH. Optimization, constraint, and history in the evolution of eyes. Q Rev Biol. 1990;65(3):281–322.

44. Piatigorsky J, Wistow G. The recruitment of crystallins: New functions precede gene duplication. Science. 1991;252(5009):1078-1080.

45. Lenski RE, Ofria C, Pennock RT, Adami C. The evolutionary origin of complex features. Nature. 2003;423(6936):139–144.

46. Lamb TD, Arendt D, Collin SP. The evolution of phototransduction and eyes. Philos Trans R Soc Lond B Biol Sci. 2009;364(1531):2791–2793.

47. Paterson S, Vogwill T, Buckling A, Benmayor R, Spiers AJ, Thomson NR, et al. Antagonistic coevolution accelerates molecular evolution. Nature. 2010;464(7286):275–278.

48. Blount ZD, Barrick JE, Davidson CJ, Lenski RE. Genomic analysis of a key innovation in an experimental *Escherichia coli* population. Nature. 2012;489(7417):513–518.

49. Meyer JR, Dobias DT, Weitz JS, Barrick JE, Quick RT, Lenski RE. Repeatability and contingency in the evolution of a key innovation in phage lambda. Science. 2012;335(6067):428–432.

50. Walker BL, Ofria C. Evolutionary potential is maximized at intermediate diversity levels. In: ALIFE 2012. MIT Press; 2012.

51. Blount ZD, Maddamsetti R, Grant NA, Ahmed ST, Jagdish T, Baxter JA, et al. Genomic and phenotypic evolution of *Escherichia coli* in a novel citrate-only resource environment. Elife. 2020;9:e55414.

52. Gould SJ. Wonderful Life. W. W. Norton; 1989.

53. Travisano M, Mongold JA, Bennett AF, Lenski RE. Experimental tests of the roles of adaptation, chance, and history in evolution. Science. 1995;267(5194):87–90.

54. Wagenaar DA, Adami C. Influence of chance, history, and adaptation on digital evolution. Artificial Life. 2004;10(2):181–190.

55. Beatty J. Replaying life's tape. J Philosophy. 2006;103(7):336–362.

56. Braught G, Dean A. The effects of learning on the roles of chance, history and adaptation in evolving neural networks. In: Australian Conference on Artificial Life. Springer; 2007.

57. Blount ZD, Borland CZ, Lenski RE. Historical contingency and the evolution of a key innovation in an experimental population of *Escherichia coli*. Proc Natl Acad Sci USA. 2008;105(23):7899–7906.

58. Flores-Moya A, Costas E, López-Rodas V. Roles of adaptation, chance and history in the evolution of the dinoflagellate *Prorocentrum triestinum*. Naturwissenschaften. 2008;95(8):697–703.

59. Keller SR, Taylor DR. History, chance and adaptation during biological invasion: Separating stochastic phenotypic evolution from response to selection. Ecol Lett. 2008;11(8):852–866.

60. Flores-Moya A, Rouco M, García-Sánchez MJ, García-Balboa C, González R, Costas E, et al. Effects of adaptation, chance, and history on the evolution of the toxic dinoflagellate

Alexandrium minutum under selection of increased temperature and acidification. Ecol Evol. 2012;2(6):1251–1259.

61. Lenski RE. Convergence and divergence in a long-term experiment with bacteria. Am Nat. 2017;190(S1):S57–68.

62. Rebolleda-Gómez M, Travisano M. Adaptation, chance, and history in experimental evolution reversals to unicellularity. Evolution. 2019;73(1):73–83.

63. Adami C, Ofria C, Collier TC. Evolution of biological complexity. PNAS. 2000;97(9):4463–4468.

64. O'Neill B. Digital evolution. PLoS Biol. 2003;1(1):e18.

65. Ofria C, Wilke CO. Avida: Evolution experiments with self-replicating computer programs. In: Artificial Life Models in Software. Springer; 2005.

66. Bundy JN, Ofria C, Lenski RE. How the footprint of history shapes the evolution of digital organisms. bioRxiv; 2021.

67. Lenski RE, Rose MR, Simpson SC, Tadler SC. Long-term experimental evolution in *Escherichia coli*. I. Adaptation and divergence during 2,000 generations. Am Nat. 1991;138(6):1315–1341.

68. Lenski RE, Wiser MJ, Ribeck N, Blount ZD, Nahum JR, Morris JJ, et al. Sustained fitness gains and variability in fitness trajectories in the long-term evolution experiment with *Escherichia coli*. Proc R Soc Lond B Biol Sci. 2015;282(1821):20152292.

69. Ostrowski EA, Ofria C, Lenski RE. Ecological specialization and adaptive decay in digital organisms. Am Nat. 2007;169(1):E1–20.

70. Brandon RN. Adaptation and Environment. Princeton Univ Press; 1990.

71. Powell R. Is convergence more than an analogy? Homoplasy and its implications for macroevolutionary predictability. Biol Philos. 2007;22(4):565–578.

72. Powell R. Contingency and convergence in macroevolution: A reply to John Beatty. J Philosophy. 2009;106(7):390–403.

73. Pennock RT. Models, simulations, instantiations, and evidence: The case of digital evolution. J Exp Theor Artif Intell. 2007;19(1):29–42.

74. Cooper TF, Ofria C. Evolution of stable ecosystems in populations of digital organisms. In: Eighth International Conference on Artificial Life. 2002.

75. Goings S, Ofria C. Ecological approaches to diversity maintenance in evolutionary algorithms. In: IEEE Symposium on Artificial Life. 2009.

76. Covert III A, McFetridge S, DeLord E. Structured populations with limited resources exhibit higher rates of complex function evolution. In: ALIFE 14: The Fourteenth International Conference on the Synthesis and Simulation of Living Systems. MIT Press; 2014.

77. Beckmann BE, McKinley PK. Evolving quorum sensing in digital organisms. In: Proceedings of the 11th Annual conference on Genetic and evolutionary computation. 2009.

78. Beckmann BE, McKinley PK. Digital Evolution of Quorum Sensing and its Applications. In: GECCO'09 Proceedings of the 11th Annual conference on Genetic and Evolutionary Computation. 2009.

79. Beckmann BE, McKinley PK, Knoester DB. Effects of communication impairments on quorum sensing. In:Third IEEE International Conference on Self-Adaptive and Self-Organizing Systems. 2009.

80. Beckmann BE, Knoester DB, Connelly BD, Waters CM, McKinley PK. Evolution of resistance to quorum quenching in digital organisms. Artif life. 2012;18(3):291–310.

81. Knoester DB, Connelly BD, Waters CM, McKinley PK. Evolution of Resistance to Quorum Quenching in Digital Organisms. Artif life. 2012;18(3):1–20.

82. Johnson AE, Strauss E, Pickett R, Adami C, Dworkin I, Goldsby HJ. More bang for your buck: Quorum-sensing capabilities improve the efficacy of suicidal altruism. arXiv preprint:14060416. 2014.

83. Lehmann KD, Goldman BW, Dworkin I, Bryson DM, Wagner AP. From cues to signals: Evolution of interspecific communication via aposematism and mimicry in a predator-prey system. PLoS One. 2014;9(3):e91783.

84. O'Donnell DR, Parigi A, Fish JA, Dworkin I, Wagner AP. The roles of standing genetic variation and evolutionary history in determining the evolvability of anti-predator strategies. PloS One. 2014;9(6):e100163.

85. Wagner AP, Zaman L, Dworkin I, Ofria C. Behavioral strategy chases promote the evolution of prey intelligence. In: Evolution in Action: Past, Present and Future. Springer; 2020.

86. Misevic D, Ofria C, Lenski RE. Sexual reproduction reshapes the genetic architecture of digital organisms. Proc R Soc Lond B Biol Sci. 2006;273(1585):457–464.

87. Weigel EG, Testa ND, Peer A, Garnett SC. Context matters: Sexual signaling loss in digital organisms. Ecology and evolution. 2015;5(17):3725–3736.

88. Chandler CH, Ofria C, Dworkin I. Runaway sexual selection leads to good genes. Evolution. 2013;67(1):110–9.

89. Canino-Koning R, Keagy J, Ofria C. Sexual selection promotes ecological speciation in digital organisms. In: Artificial Life Conference Proceedings 14. MIT Press; 2017.

90. Wiser MJ, Ribeck N, Lenski RE. Long-term dynamics of adaptation in asexual populations. Science. 2013;342(6164):1364–1367.

91. Wiser MJ, Dolson EL, Vostinar A, Lenski RE, Ofria C. The boundedness illusion: Asymptotic projections from early evolution underestimate evolutionary potential. PeerJ Preprints. 2018;6:e27246v2.

92. Devosoft. Avida. Digital Evolution Software; 2021.Available from: https://avida.devosoft.org/

93. Sokal RR, Rohlf J. Biometry. W. H. Freeman; 1994.

94. Bundy JN. Repository: History and habitat novelty shape pleiotropy. GitHub; 2021. Available from: https://github.com/bundyjay/Avida_footprint_pleiotropy_gof

CHAPTER 4: EXPLORING THE EFFECTS OF HISTORY ON EVOLUTION: AN AVIDA-ED EXERCISE

Authors: Jason Nyerere Bundy, Zachary Blount, and Richard E. Lenski

Abstract

In *Wonderful Life*, paleontologist Stephen Jay Gould proposed the thought experiment of "replaying life's tape" to highlight the importance of historical contingency. In essence, Gould thought that a lineage's history—including seemingly innocuous, chance events without obvious importance when they occurred—could greatly influence later evolution. This thought experiment has become a framework for experimental tests of contingency in evolution, typically called replay experiments. We designed an Avida-ED exercise based on a two-phase replay experiment that allows students to investigate how the unique evolutionary history of populations that evolved under identical conditions during the first phase can influence the subsequent evolution of descendant populations in a new environment during the second phase. The exercise is intended as an advanced supplement to the core exercises in the Avida-ED manual. However, it could also be used by students familiar with the basics of Avida-ED and the fundamental principles of evolution. Before tackling this exercise, students should be familiar with the concepts of random mutation, asexual inheritance, and adaptation via natural selection. Completing this activity will help students understand how various evolutionary mechanisms interact to determine the results of the evolutionary process.

Introduction

We have designed an exercise to explore the effects of history on evolution using the Avida-ED educational platform. The activity features a two-phase replay experiment with an experimental design similar in structure to those used by evolutionary biologists working with microbes. In the first phase, we generated lineages with unique histories by independently evolving 15 replicate populations from a common ancestor under identical conditions. The design of the first phase is similar to Lenski's Long-Term Evolution Experiment (LTEE) with *E. coli* that features 12 evolving populations of bacteria derived from a common ancestor in a simple environment where glucose is the primary food source [1,2]. As an intermediate step, we screened a single evolved genotype from each population in the alternative environment that will be used for the second phase. We identified three cases where the organism's prior history resulted in unique evolvability in the new environment, which rewarded certain additional functions. We saved the selected genotype from these three lineages to serve as common ancestors for replicate populations that the students will evolve in that same new environment during the second phase. The second phase is similar to the second phase of Travisano et al.'s experiment that used clones derived from a single genotype from each LTEE population as ancestors for replicate populations that evolved in a new environment where maltose replaced the glucose used in the original LTEE environment [3]. Following the second phase, students record each population's average fitness and whether the newly rewarded functions have evolved. The resulting data are used to test whether the unique *history* that each ancestor experienced in one environment during the first phase affects the evolution of their descendants in a new environment during the second phase.

Background

In his book *Wonderful Life*, the late paleontologist Stephen Jay Gould proposed the thought experiment of "replaying life's tape" to better understand the predictability of biological evolution [4]. His now-famous thought experiment has become a widely used experimental framework for studying the role of historical contingency in evolution. Of course, Gould's vision of replaying the history of life on Earth isn't feasible. However, he used the imagined results of that experiment to advance an unpredictable, contingent view of nature—one in which past events, including seemingly innocuous chance occurrences, could greatly alter the evolutionary trajectory of life on Earth. In Gould's view, any repetition of evolution from, say, the time of the Burgess Shale (~500 million years ago, in the middle of the Cambrian Period) would produce a very different and unrecognizable biota. By contrast, some others, including another paleontologist, Simon Conway Morris, have argued that the prevalence of convergent evolution—in which similar traits evolve independently, often even in distantly related lineages—demonstrates that the limited number of viable solutions to any particular environmental challenge, coupled with the power of natural selection, makes evolution predictable [5]. Consequently, replaying evolution should generate an alternative history that is strikingly similar to the actual history of life on Earth. Even though Gould's thought experiment remains out of reach on the grand scale that he imagined, it has stimulated research into the potentially idiosyncratic nature of history in evolution [6].

In particular, a burgeoning research program has emerged that uses Gould's thought experiment as a framework to investigate the role of contingency in evolution. This work includes both comparative studies that analyze diverse lineages as though they were natural replay experiments and laboratory experiments using rapidly evolving microbes and digital organisms. Some of the experimental investigations have sought to study the role of history within the context

of evolutionary responses to environmental change. These "historical difference" replay experiments involve two phases. In the type of experiment implemented here, the first phase is performed to generate unique evolutionary histories by allowing replicate populations to evolve from a common ancestor in identical environments. These replicate populations represent simultaneous, rather than sequential, replays of evolution from the same starting point and under identical conditions. Although the populations start out identical, over time the interaction of stochastic processes, including the identity and order of mutations as well as genetic drift, and the more deterministic process of natural selection, cause the populations to diverge from one another. A single genotype is chosen from several lineages that evolve during the first phase. These genotypes are then used to initiate many replicate populations that evolve under identical conditions, but in an environment that differs from the environment used during the first phase. The second phase, like the first, employs simultaneous replays of evolution under identical conditions. However, during the second phase, each of several groups of replicates start from a different progenitor genotype that had a unique, independent history created during the first phase. Any systematic differences in trait evolution between these groups after the second phase implies a lingering effect of their historical differences generated during the first phase.

In sum, Gould's thought experiment of "replaying life's tape" is essential to both the philosophical and methodological foundations of this exercise, as it informs both why and how we study the effects of history on evolution.

Developing the Exercise

We first evolved 15 populations in Avida-ED. Each population was founded by the default ancestor, which can reproduce but not perform any of the nine computational tasks that are potentially awarded in Avida-ED. These populations evolved for 10,000 updates (the unit of time

in Avida-ED) in an environment with three computational resources available: notose, nanose, and andose. Avidians can obtain the rewards associated with these resources if they perform the logic functions NOT, NAND, and AND, respectively. The first two are relatively easy functions to evolve, while the third is of medium difficulty. The genome of the default ancestor contains a copy loop, which allows it to copy its genome, but it is otherwise filled with non-functional instructions (nop-C). As the organisms copy their genomes, point mutations change one instruction to another following a mutation rate set by the experimenter. We set the rate to 5% per site, which corresponds to an average of 2.5 mutations per offspring given the fixed genome length of 50 instructions. After 10,000 updates we chose a single organism from each population that served as a new ancestor for a set of 30 replicate populations, which then evolved in the second phase environment for 2,000 updates. This initial second phase was performed to identify potential founders for the exercise that demonstrated unique evolutionary potential in the new environment.

In this second phase, all populations evolved in an environment that included two additional resources, orose and norose. Avidians can gain energy from these resources when they perform the logic functions OR and NOR, which are of medium and high difficulty, respectively. Owing to technical difficulties with Avida-ED, we were unable to collect data for two of the 15 ancestors in the second phase. For the remaining 390 populations (13 ancestors x 30 replicates), we assessed the final average fitness for each population and whether organisms evolved the ability to use the two new resources, orose and norose.

We used these data to identify three ancestors that appeared to show distinct outcomes during the second phase, and which might therefore be used in the classroom exercise. We then performed three additional rounds of the second phase, each time evolving 50 replicate populations from each of the three candidate ancestors to ensure that groups of replicates derived from these

founders produced consistent results. We named these organisms ancestors A, B, and C. Ancestor B typically evolves the ability to use both new resources, thereby achieving high fitness; Ancestor C rarely evolves either ability and thus attains low fitness in the new environment; and Ancestor A tends to be intermediate. Since we selected these founders from populations that evolved from an identical genotype under identical conditions during the first phase, the distinctive tendencies between groups of replicates derived from these ancestors during the second phase reflect their unique histories that arose during the first phase in the ancestral habitat. We chose these ancestors to ensure students will observe unique tendencies between groups.

Following an introduction to the exercise, students are given several conceptual questions and asked to predict what will happen in the second phase of the experiment. Students are then divided into three teams, and each team is assigned one ancestor. Each student evolves 10 replicate populations ("replays") from their assigned ancestor in the new environment for 2,000 updates. After the runs are complete, students work with their team to assess the overall mean fitness and proportion of populations that evolved to use the new resources across all replicates from the ancestor assigned to their team. Students then share this information among the three teams to assess and compare the overall outcomes for the replicate populations derived from the three different ancestors. Finally, the students should revisit their predictions before answering a concluding set of conceptual questions.

Although the exercise was designed for a traditional classroom setting, a slightly modified version is available for remote instruction. The exercise can also be split into two class periods; alternatively, a portion of the exercise can be done as homework.

After we had conceptualized the activity, identified ancestors, and collected sufficient data to be confident about the consistency of the experimental results, we then drafted a handout with

the instructions necessary to lead students through the activity. We included background information to introduce students to the concept of historical contingency and to Gould's thought experiment of "replaying life's tape." We included a link to a video segment from *Through the Wormhole with Morgan Freeman* that features Richard Lenski's Long-Term Evolution Experiment with *E. coli* [7], so that the students can see an example of a celebrated replay experiment in a biological setting. We also provide two suggested readings for advanced or especially motivated students who want to learn more about the evidence for historical contingency in evolution.

We tested the timing of the exercise and data-collection protocols during a guest lecture by author J.B. to a class of undergraduate students in the Lyman Briggs College at Michigan State University. Following this trial run, we made some modifications to the handout for students. We also prepared instructions to assist instructors preparing to use the activity with their classes. These instructions include additional background information and suggested readings. In addition, author J.B. recorded a video to demonstrate the exercise in conjunction with the Avida-ED team and WKAR Public Media at Michigan State University.

APPENDIX

APPENDIX

Instructor Handout: Exploring the Effects of History on Evolution

Jason N. Bundy, Zachary D. Blount, and Richard E. Lenski

Background Information

This exercise examines the concept of historical contingency. In brief, that is the idea that a lineage's history, including even chance events, can strongly influence its later evolution. Students will perform a variation on Gould's thought experiment of "replaying life's tape" to investigate how adaptation to a new environment can be influenced by random events that occurred in the past, and which led to differences among the three ancestors that the students will use in their experiments. This exercise is appropriate for advanced students who already understand the core concepts of random mutation and natural selection. It might be appropriate, therefore, for an AP or other advanced high school biology class, an upper-level undergraduate course, or even a graduate course. The students will benefit from a basic familiarity with Avida-ED and/or having completed the primary exercises in the Avida-ED lab manual.

Abstract

This exercise provides an advanced supplement to the exercises in the Avida-ED lab manual. It builds upon students' understanding of core evolutionary concepts including random mutation, asexual inheritance, and adaptation by natural selection. The focus of this exercise is the unpredictable, contingent nature of evolution. The students will conduct the second phase of a two-phase experiment. In the first phase, which we have completed, three independent populations evolved from a common ancestor in identical environments. We then chose one individual (clone) from each first-phase population to serve as founding ancestors for the second phase. In the second phase, three teams of students will evolve many replicate populations from one of the founding ancestors in a new environment that favors certain functions that were not rewarded in the first-phase environment. Students will examine how the unique histories of the founding ancestors affect the potential for their descendants to adapt to the new environment. Instructors will guide students in starting the experiments, collecting data, and discussing open-ended questions.

The Big Ideas

In a widely read book, the paleontologist Stephen Jay Gould proposed the thought experiment of "replaying life's tape" in order to emphasize the importance of historical contingency in evolution. Gould argued that evolution is inherently unpredictable, such that any repetition of the process would result in a different outcome, with an array of unrecognizable life forms. However, some others, including the paleontologist Simon Conway Morris, argued that instances of convergent evolution—in which similar traits, such as wings and eyes, have evolved in distantly related lineages—demonstrate that the power of natural selection makes evolution predictable, at least in broad outline. According to this line of thought, life forms would be basically familiar even if evolution was somehow replayed. While the idea of replaying evolution on the scale imagined by Gould is fanciful, researchers today are performing conceptually similar experiments to examine contingency in evolution on a smaller scale.

In particular, researchers are often interested in studying the role of contingency when populations encounter new environments. How does adaptation to one environment promote or

constrain the potential for descendants to adapt to a new environment? This exercise features a two-phase evolution experiment that allows students to explore how this famous thought experiment has been transformed into an experimental approach that can be used to study the role of history in evolution.

Additional Readings

- Stephen Jay Gould. 1989. *Wonderful Life*. Norton, New York, NY.

Gould was a paleontologist, evolutionary biologist, historian of science, and a popular writer on evolution for the general public. Much of his thinking about evolution revolved around the idea that evolutionary outcomes depend on the particular, unique, and even quirky events in history—a notion that is often called historical contingency. *Wonderful Life* describes this idea in the context of an important fossil deposit in Canada called the Burgess Shale. The Burgess Shale contains the remains of early animals from about 500 million years ago, many of which appear bizarre and have unknown affinities to the animals of today. Gould thought that an important lesson from those fossils is that life could have evolved along very different paths, had slightly different events occurred—even differences so small that they would have seemed inconsequential when the events took place. He illustrated his view with the thought experiment of "replaying life's tape". The design of this exercise is inspired by Gould's thought experiment.

- **Publisher's page**: https://www.norton.com/books/Wonderful-Life/

- Simon Conway Morris. 2003. *Life's Solution.* Cambridge University Press, Cambridge, UK.

Conway Morris is a paleontologist who takes a very different view than Gould. He argues that the vagaries of historical contingency did not much matter because there are a limited number of viable paths for evolution to follow. If we could replay the history of life on Earth, he argues that one would see a living world that is very similar to the one in which we live. As evidence, he points to the pervasiveness of convergent evolution—that is, lineages that independently evolved similar features, such as wings and eyes. Much of this book is devoted to fascinating descriptions of many instances in which different lineages evolved very similar traits as adaptations to similar conditions. Ironically, Conway Morris studied the Burgess Shale fossils, and he is the central character in Gould's *Wonderful Life*, even though they came to opposing positions on the importance of historical contingency in evolution!

- **Publisher's page**: https://doi.org/10.1017/CBO9780511535499

Objectives for Student Learning
During this lesson, students will:

- Investigate how researchers implement Gould's thought experiment of "replaying life's tape" in an evolution experiment.
- Execute a unique experimental design for testing evolutionary mechanisms.
- Analyze the results of evolutionary "replays" to decide if adaptation to one environment can be influenced by a lineage's history in some previous environment.
- Create data to evaluate whether random variation among independent replicate populations (replays) can lead to systematic differences between their descendant populations.

Students' Prior Knowledge
Most students have experience thinking about chance encounters and random events that have the potential to change their lives. Some will likely have read or heard about how such seemingly small events can even potentially alter the course of history. The numerous ways in which chance events that occurred in the past can influence later events will provide a basis for students to think about the role of such contingencies in the evolutionary history of life on Earth.
Students also often think about their own abilities in a way that is context dependent. Students may be stronger in certain subjects, or they may find particular games and activities easier than alternatives. These differences in individual strengths may be useful for helping students think about how the performance of an individual (or species) depends on the situation, and which capabilities are more important in a particular environment. Even though evolutionary theory is about change in populations (rather than the development of individuals), it can be helpful for students to think about how their unique traits have a distinct evolutionary history.
Students also have experience discussing how the culture in which they live is different than that of their parents, grandparents, and earlier generations of their family or community. This can help students think about how environments change over time, throughout the course of a lineage's history. Our modern lives involve new technologies, and thus we need different skills and resources than our ancestors who lived long ago, although many of our traits arose long before our species met our current habitats.

Conceptual Connections

Biological Connections	Nature of Science Connections
Theory of evolution Experimental evolution	Science practices (experimentation) Collaboration and peer review

Materials
Student Handout: "Exploring the Effects of History on Evolution"
EEHE_experiment_setup.zip (contains student handout, completed examples, and setup files)

Lesson Plan

Introduction
Interactive introductory lecture and discussion. Review key ideas as necessary (see objectives).

Establish the scenario. Explain to students that they will be conducting the second part of a two-phase experiment. The first phase has already been performed by an Avida-ED research team at Michigan State University. To help the students understand the experiment, lead the class through their reading of the exercise, up until the optional video. Whether or not you decide to show the video, have your students read the description that briefly introduces Richard Lenski's *E. coli* Long-Term Evolution Experiment. Continue reading until you reach the start of Part 1 on page 4.

Suggested Video (~5 minutes, described in student handout)
https://www.sciencechannel.com/tv-shows/through-the-wormhole/videos/evolution-is-like-poker

Main Activities (~2 hours)

Part 1: Understanding the previously generated Ancestors with their unique evolutionary histories (15 minutes). The students should read Part 1 of the activity and answer the questions.

Part 2: Performing the experiment (60 minutes): Students use Avida-ED to perform the experimental runs and collect their data according to the instructions for Part 2. **Divide students into three teams: A, B, and C before you begin.** They will perform their runs using the workspace file in the materials folder. Make sure each student has access to this file before conducting Part 2.

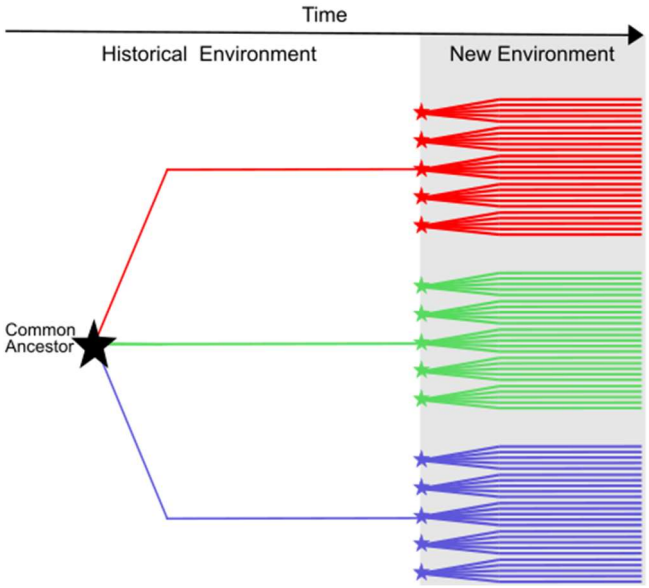
Part 3: Working with the group (15 minutes): During this period, students work in one of three groups to combine the data from all of their replicate runs that started from a particular ancestor. At least two students should record and check the data in each group.

Part 4: Recording data from the entire class and discussion (30 minutes): Each group should choose a representative to report their data to the entire class. Each group will then record the data from the entire class by filling out the chart in Part 4. Students will then work together to answer the Discussion Questions and Wrap-up.

Note: These activities could be broken up in various ways, depending on the course schedule and available time. For example, the Introduction, optional video, and Part 1 might all be done in one in-class period; Part 2 could be performed as homework; and Parts 3 and 4 completed in a second in-class period.

Student Handout: Exploring the Effects of History on Evolution

Jason N. Bundy, Zachary D. Blount, and Richard E. Lenski



- In biology, populations generate variation by mutations that occur at random.
- Populations that start out identical will accumulate different mutations in different orders.
- Even beneficial mutations are often lost by random genetic drift while they are still rare.
- Because of the interaction between the random processes of mutation and genetic drift, and the nonrandom effects of natural selection, each lineage experiences a unique history.
- How does a lineage's unique history affect its potential to adapt to a new environment?

Student Learning Goals

- Students will investigate the repeatability of evolution by performing a particular version of a replay experiment, inspired by paleontologist Stephen Jay Gould's thought experiment of "replaying life's tape".
- Students will analyze how a lineage's history can affect its likelihood of evolving a new trait.
- Students will evaluate how the random processes of mutation and genetic drift interact with the deterministic process of natural selection during evolution.

Driving Questions

- Is evolution predictable?
- How can we understand and explain differences between replicate populations that begin with the same ancestor and evolve under identical conditions?
- Can differences between ancestors that evolved in the same historical environment affect the ability of their descendants to adapt to a new environment.

The Role of History in Evolution

We've all heard stories about how chance encounters can profoundly affect lives. Maybe it's a story about how your parents or grandparents met unexpectedly at some event, or when someone's car broke down. Without that chance meeting, you'd probably not even exist today. Shakespeare used the line "What's past is prologue" in *The Tempest* to make the point that the past sets the context in which the present takes place. The past is always with us, so that what happens today depends on everything that came before.

History also plays a central role in the science of evolution. All life on Earth shares a common history. There might be other planets with life, but there is only one Earth, and—unless there are parallel universes—life on Earth has a particular and unique history. Nonetheless, one can try to imagine how life today might have been different had certain events not happened in that history. (Students sometimes think that science is all about existing facts. But imagining, and trying to predict, what would happen under different scenarios is a key part of the scientific mindset.) If evolution were to happen all over again, would life on Earth be the same as we see it today? Probably not, but just how different life would be is a matter of scientific interest and debate. Would multicellular organisms have evolved, no matter what? Would the land have eventually been colonized by fish? Would bipedal organisms with unusually large brains have inevitably evolved? Of course, we can never know the answers to these particular questions without a time machine or replicate planet Earths. But researchers can examine the more general question of whether the *process* of evolution is repeatable. Evolutionary biologists have addressed this question by performing experiments with fast-reproducing organisms (such as bacteria) in which they observe many evolutionary histories play out under identical conditions, and then measure the outcomes. These experiments allow researchers to "replay life's tape" in the lab.

Replaying life's tape: Is evolution repeatable or unpredictable?

Before starting a replay experiment, researchers isolate and make several copies (often called *clones*) of a single organism, which they call the *common ancestor*. They then start a separate population with each copy, which they call a *replicate*. Because each replicate population started from an identical copy of the common ancestor, researchers can be sure that they all began evolving from the same starting point. The researchers then allow the replicate populations to evolve under identical environmental conditions.

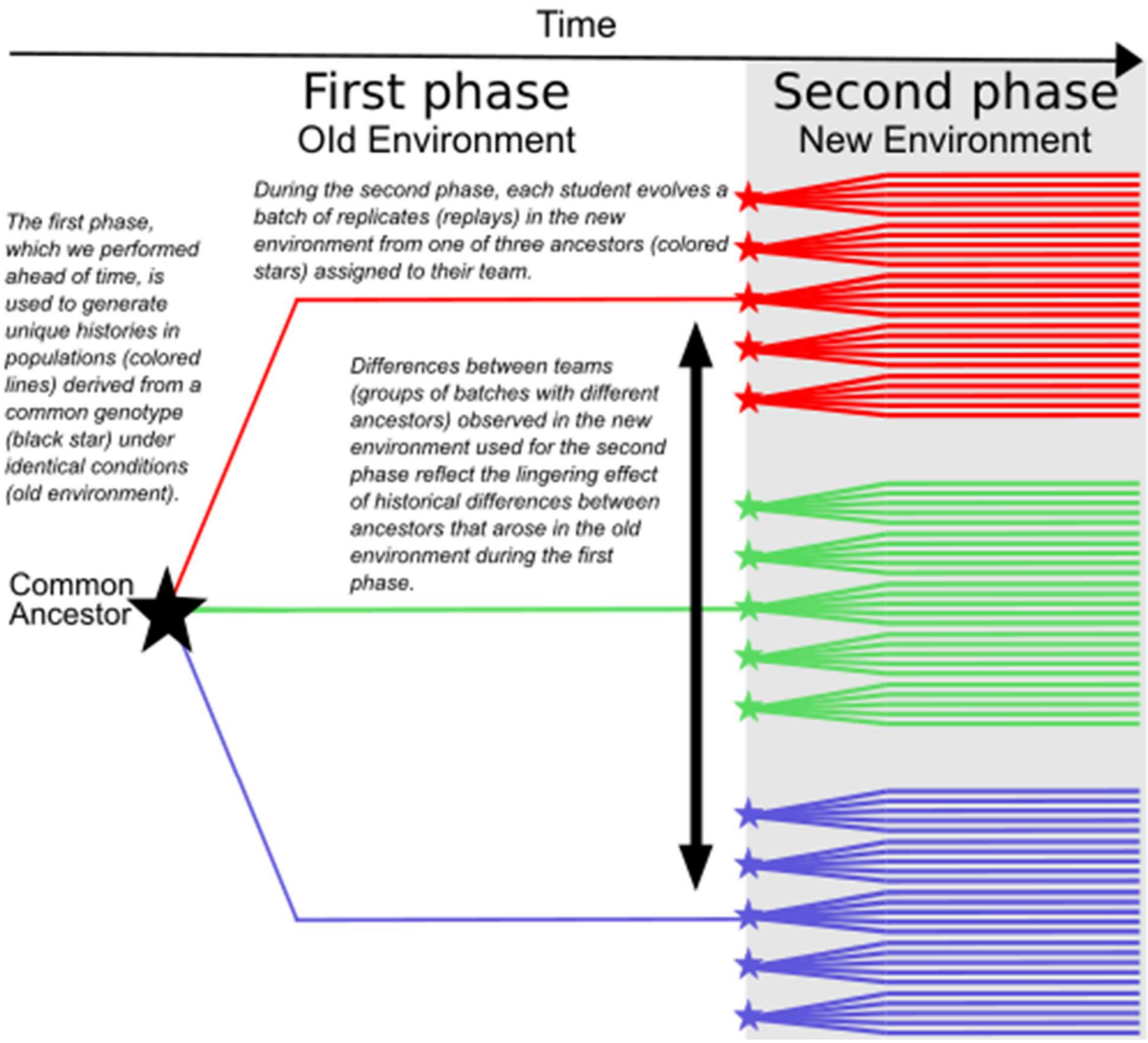
Even though the replicate populations in a replay experiment start off identical and evolve under identical conditions, they begin to differ as they evolve. Different mutations will occur at random in each population at different times. Some of those mutations will be lost at random by genetic drift. At the same time, the populations will systematically adapt to the environment, as natural selection tends to increase the frequency of any beneficial mutations that arise and eliminate those mutations that harm the organisms' performance. As a consequence, each population will tend to get better at growing in the environment over time. But which mutations are favored and spread will likely differ between the replicate populations, depending on what mutations occur and when. This interaction between the random processes of mutation and genetic drift and the systematic, non-random process of natural selection will cause each population to experience its own unique *history*. Those different histories matter because the effects of later mutations often depend on the mutations that accumulated during a lineage's history. This sort of experiment is like replaying the "tape of life" from the same point many times simultaneously. Researchers can then observe and measure the similarities and differences that occur over time.

Suggested activity: Watch a short video
https://www.sciencechannel.com/tv-shows/through-the-wormhole/videos/evolution-is-like-poker

This video is from the television show called *Through the Wormhole with Morgan Freeman*. In this episode, evolutionary biologist Richard Lenski discusses a long-term experiment with 12 evolving populations of *E. coli* bacteria that have been growing in the lab since 1988. All of the populations had the same ancestor and have evolved in the same environment, which contains two substances the bacteria could possibly eat: glucose and citrate. Glucose is the preferred food for *E. coli*, and all of the populations grew (and continue to grow) on it. Most *E. coli*, including the experiment's ancestor, cannot grow on citrate. However, after about 15 years, one lineage evolved the ability to grow on the citrate, which allowed that population to get several times larger. Even after 19 more years, no other population has evolved that ability. Lenski explains that the random production of mutations enabled this one population to solve the problem of how to grow on citrate, and he draws an analogy with the game of poker, where the value of a randomly drawn card (a new mutation) depends on the other cards already in the player's hand (other mutations that came before in a lineage's history).

How does a lineage's history influence its future in a new environment?
In nature, environments rarely stay the same forever. Changing environmental conditions create opportunities for organisms to evolve features that are different from those of their ancestors, and which adapt them to their new habitat. For example, mammals like dolphins and whales, which have evolved in and adapted to aquatic environments, have different features than their ancestors, which walked on land. When organisms adapt to an environment that is different from their ancestor's habitat, biologists want to understand how a lineage's *history* in the old environment affects its potential adaptation to the new environment.

A slightly different type of replay experiment can be performed to investigate this question. To do so, researchers add a *second phase* to the experiment. For this second phase, researchers again isolate individual organisms that are used to start new replicate populations. However, in this case, they isolate an individual from each of two or more different populations at the end of the first phase of the experiment. These individuals thus have unique evolutionary histories that occurred in the old environment, as illustrated in the figure below. The researchers then start new replicate populations with identical copies (clones) of individuals that evolved in the *historical* environment. They then let these populations evolve in and adapt to a *new* environment. Because of the multiple repetitions of evolution that started from each historical lineage from the first phase of the experiment, it is possible to examine whether and how a particular ancestor and its unique history influences adaptation to the new environment during the second phase.


Time
First phase
Old Environment
Second phase
New Environment
The first phase, which we performed ahead of time, is used to generate unique histories in populations (colored lines) derived from a common genotype (black star) under identical conditions (old environment).
During the second phase, each student evolves a batch of replicates (replays) in the new environment from one of three ancestors (colored stars) assigned to their team.
Differences between teams (groups of batches with different ancestors) observed in the new environment used for the second phase reflect the lingering effect of historical differences between ancestors that arose in the old environment during the first phase.
Common Ancestor

Part 1: Generating unique evolutionary histories

For convenience, we already performed the first phase of this experiment. We started three populations from the default Avida-ED ancestor, which can reproduce but cannot perform any other functions. We evolved these populations for 10,000 updates in an environment with only three resources. Two of the resources—notose and nanose—are easy for Avidians to evolve the ability to consume, while the other, andose, is a bit more difficult. After the 10,000 updates, we chose one of the individual Avidians with the highest fitness from each population. Each of them will serve as new ancestors for the second phase of the experiment, which you will perform in a new environment with two additional resources: orose and norose. The table below shows the fitness of each ancestor in the historical environment, along with the resource-acquisition functions that it can perform.

Ancestor	Fitness	Functions Performed
A	1.34	NOT, NAN
B	3.19	NOT, NAN, AND
C	1.45	NOT, NAN

You will perform the second phase of the experiment. You will evolve replicate populations started from each ancestor in the new environment for 2000 updates.

Before You Begin the Experiment, Answer the Following Questions:

- After 2000 updates, which ancestor's descendants do you think will have the highest average fitness in the *new* environment? Which ancestor's descendants do you think will have the lowest average fitness in the new environment? Explain your thinking.
- Do you think a population's ability to evolve to consume the new resources (orose and norose) will depend on which ancestor it evolved from? Explain your thinking.
- If the three ancestors (A, B, and C) all started from the same common ancestor at the start of the first phase, and if they all evolved under identical conditions, how can we explain the differences in their fitness and ability to consume the andose resource?

Part 2: Performing the Experiment

1. Open Avida-ED (https://avida-ed.msu.edu/app4/).
2. Go to file, Open Workspace, and open the workspace file (e.g.,"workspace.zip").
3. In the "Configured Dishes", drag the appropriate dish (e.g., "Team_A_Dish") onto the "@default" textbox located next to the small petri dish over the top left-hand side of the population viewer. The title bar for the configured dish will turn from red to green when you hover over the correct location. After completing this task, the "@default" text should be replaced with the name of your configured dish (e.g., "Team_A_Dish").
4. Flip to the "Setup" screen and check the following parameters:
    - Dish Size: 30 x 30
    - Per Site Mutation Rate: 5%
    - Ancestral Organism: Ancestor_A, Ancestor_B, or Ancestor_C (depending on your assigned group)
    - Place Offspring: Near their parent
    - Resources (unlimited): notose, nanose, andose, orose, norose
    - Repeatability Mode: Experimental
5. Flip back to the "Stats" screen. Check that "Pause Run at Update" is set to 2000 and that the box is checked.
6. Open nine additional browser tabs and repeat steps 1-5 in each tab. This will allow you to perform 10 runs at once. (If you have an old computer or a slow internet connection, you may want to perform only five runs at a time. However, most users should be able to perform 10 runs at once.)
7. Press "run" at the bottom of the population viewer in each tab and allow the experiment to run. You can start recording your data when each run finishes. **After update 2000,** record the average fitness of the population and whether any organisms evolved to perform ORO (consume orose) or NOR (consume norose) in the table below.

Rep	Fitness	ORO (y/n)	NOR (y/n)		Rep	Fitness	ORO (y/n)	NOR (y/n)
1					6			
2					7			
3					8			
4					9			
5					10			

8. Use the table below to record the number of repetitions that you ran. (This should be 10, unless your instructor told you to run a different number.) Then record the average final fitness of your populations, as well as the number of populations that evolved the ability to perform the ORO and NOR functions.

Number of Repetitions	
Average Fitness	
Number of Populations that Perform ORO	
Number of Populations that Perform NOR	

After you record your data, you have completed the individual portion of this exercise. You will do the rest of this exercise with your group and classmates.

Part 3: Collecting Group Data

1. For each student in your group, use the table below to record the average fitness of the 10 repetitions and the number of populations (out of 10) that evolved to perform the "ORO" and "NOR" functions. (You will need more than one sheet if your group has more than 10 members.) At least two people from your group should independently record these data to ensure accuracy.

Enter group data on the shared table below.

Ancestor (circle): **A** **B** **C**

Student	Average Fitness	Number of ORO Populations	Number of NOR Populations		Student	Average Fitness	Number of ORO Populations	Number of NOR Populations
1					6			
2					7			
3					8			
4					9			
5					10			

2. Calculate the overall mean fitness for your group by summing the average fitness for all members and dividing by the number of students in your group who provided data. Also, record the total number of populations (repetitions) run by your entire group and the total number of populations that evolved the ability to perform each new function. Again, make sure at least two members of your group independently calculate and record this information, and resolve any discrepancies.

Enter group average and total number of populations that utilize each new resource.

Ancestor (circle): **A** **B** **C**

Group Average Fitness		Total Number of ORO Populations	
Total Number of Populations		Total Number of NOR Populations	

Part 4: Recording Data from the Entire Class

1. Each group should nominate a reporter who will share the group's data with the rest of the class.
2. Each group should complete the table below with the data from the entire class. For each of the three different ancestors (A, B, and C), calculate and record the overall average fitness, the total number of populations, and the total number of populations that evolved to use each new resource.

Enter group averages and total number of populations that can use each new resource.

Ancestor	A	B	C
Average Fitness			
Total number of populations			
Total Number of ORO Populations			
Total Number of NOR Populations			

Discussion Questions and Wrap-up. After examining the class data, discuss with your lab team and, together, respond to the following questions.

➔ Revisit the section titled "Before You Begin the Experiment, Answer the Following Questions." Are the results that your class observed consistent with your predictions? Which predictions were correct? Which ones were wrong? (There's nothing wrong with incorrect predictions – that's often how science makes progress!)

➔ How does the unique history of every lineage result from the interaction between the random effects of mutation and genetic drift, on the one hand, and the process of adaptation by natural selection, on the other hand?

➔ How did your class use replay experiments to investigate the effect of evolutionary history in one environment on later evolution in a different environment?

➔ Stephen Jay Gould said, "replay the tape [of life] a million times … and I doubt that anything like *Homo sapiens* would ever evolve again." What do you think he meant when he said this considering the results you saw during this activity?

Suggested Readings for Interested Students

- Zachary D. Blount. Replaying Evolution. *American Scientist* 105, 156-165 (2017).

In this article from a popular-science magazine, Blount (who did much of the work on the evolution of citrate described in the introduction to this exercise) discusses history's role in evolution and the experiments that biologists have performed to understand it. The overarching question is the extent to which evolution is repeatable, and what we have learned from the work of many researchers interested in that question.

- **Link**: https://tinyurl.com/2wrpmevt

- Zachary D. Blount, Richard E. Lenski, Jonathan B. Losos. Contingency and determinism in evolution: replaying life's tape. *Science* 362, eaam5979 (2018).

This review article describes the debate in biology over the role of history in evolution, and whether outcomes are contingent and unpredictable, as the late Stephen Jay Gould thought, or highly repeatable, as the paleontologist Simon Conway Morris thinks. It also describes the various approaches that scientists use to examine history's role and to investigate the repeatability of evolution. This article provides an in-depth overview of the available evidence concerning these issues, along with related questions that would benefit from further research.

- **Link**: https://tinyurl.com/5e5hsysk

REFERENCES

REFERENCES

1. Lenski RE, Travisano M. Dynamics of adaptation and diversification: a 10,000-generation experiment with bacterial populations. Proc Natl Acad Sci USA. 1994;91(15):6808–6814.

2. Lenski RE. Convergence and divergence in a long-term experiment with bacteria. Am Nat. 2017;190(S1):S57–68.

3. Travisano M, Mongold JA, Bennett AF, Lenski RE. Experimental tests of the roles of adaptation, chance, and history in evolution. Science. 1995;267(5194):87–90.

4. Gould SJ. Wonderful Life. W. W. Norton; 1989.

5. Morris SC. Life's Solution. Cambridge Univ Press; 2003.

6. Blount ZD, Lenski RE, Losos JB. Contingency and determinism in evolution: Replaying life's tape. Science. 2018;362(6415):1–10.

7. Evolution is Like Poker- Through the Wormhole with Morgan Freeman. 2015 Available from: https://www.sciencechannel.com/tv-shows/through-the-wormhole/videos/evolution-is-like-poker

8. Blount ZD, Borland CZ, Lenski RE. Historical contingency and the evolution of a key innovation in an experimental population of *Escherichia coli*. Proc Natl Acad Sci USA. 2008;105(23):7899–7906.

9. Blount ZD. Replaying evolution. Am Sci. 2017;105(3):156–165.