This is to certify that the
dissertation entitled

DIGITAL SEX: CAUSES AND CONSEQUENCES OF RECOMBINATION

presented by

DUSAN MISEVIC

has been accepted towards fulfillment
of the requirements for the

Ph.D. degree in Zoology
Ecology, Evolutionary Biology,
and Behavior Program

Richard E. Lenski
Major Professor's Signature

May 8, 2006
Date

DIGITAL SEX: CAUSES AND CONSEQUENCES OF RECOMBINATION

By

Dusan Misevic

A DISSERTATION

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

DOCTOR OF PHILOSOPHY

Department of Zoology
Ecology, Evolutionary Biology, and Behavior Program

2006

ABSTRACT

DIGITAL SEX: CAUSES AND CONSEQUENCES OF RECOMBINATION

By

Dusan Misevic

The evolution of sexual reproduction has long been a major problem in biology. It is a vibrant and fascinating area of research, rich with theories but poor in experimental results. The crux of the sexual paradox is the dominance of the sexual mode of reproduction in the natural world in spite of the apparent costs associated with recombination. In this dissertation I test two different theories about the evolution of sexual reproduction and also investigate effects of recombination on genetic architecture.

According to one theory, sex opposes the fitness-destroying process of Muller's ratchet, which occurs by the stochastic loss of high-fitness genotypes in small populations. Sex opposes the ratchet by allowing genotypes with different deleterious mutations to produce mutation-free offspring. In Chapter 1 I used the Avida digital-evolution software to investigate sex in relation to Muller's ratchet. Populations of digital organisms mutated, competed, and evolved in a complex environment. Populations were either asexual or sexual; in the latter case, parental genomes recombined to produce offspring. I also varied genomic mutation rates and population sizes, which at extreme values often caused mutational meltdowns and population extinctions. My results demonstrate that sex is advantageous for population survival under some conditions. However, differences in extinction probabilities were usually small, occurred over a narrow range of mutation rates and population sizes, and the mean fitness of surviving asexual populations was often greater than in sexual populations.

Theory suggests that modularity promotes evolvability, and that aggravating (synergistic) epistasis among deleterious mutations facilitates the evolution of sex. In Chapter 2, by contrast, I investigate how genetic architecture is shaped by reproductive mode. I allowed 200 populations of digital organisms to evolve for over 10,000 generations while reproducing either asexually or sexually. For ten randomly chosen organisms from each population, I constructed and analyzed all possible single mutants as well as one million mutants at each mutational distance from two to ten. The genomes of sexual organisms were more modular than asexual ones; sites encoding different functional traits had less overlap and sites encoding a particular trait were more tightly clustered. Net directional epistasis was alleviating (antagonistic) in both groups, although the overall strength of this epistasis was weaker in sexual than asexual organisms. My results show that sexual reproduction profoundly influences the evolution of the genetic architecture.

In Chapter 3 I demonstrate that sex can evolve *de novo* and outcompete the asexual mode of reproduction under changing environmental conditions. I evolved large populations of digital organisms for thousands of generations in six environments with different periods of substrate change. Sex was the dominant mode of reproduction when the environments were changing rapidly, with up to 65% of populations evolving to reproduce sexually. The ancestral reproductive mode and genetic architecture had only weak effects on whether populations evolved to be sexual or asexual. I also found that in the environmental conditions where sex was prevalent, the sexual populations on average had higher fitness than asexual ones. Chapter 3 experimentally demonstrates the importance of changing environments for the evolution of sex.

ACKNOWLEDGEMENTS

The road to my Ph.D. degree has definitely been long and winding, but fortunately has been lined by a great number of colleagues, collaborators, and friends. While in more ways than one I typically avoided a straight and narrow path, they helped keep me on track, and I certainly would have never made it without their support, advice, and help.

First and foremost, I must acknowledge and deeply thank my advisor, Richard. E. Lenski. For years I have been jokingly proclaiming that "I want to be a scientist when I grow up" and Rich has been instrumental in getting me much closer to that goal. I look forward to conducting independent research in the future, but sometimes wish I could spend 6 more years learning from and working with Rich. He is a truly incredible mentor and scientist, and I can just hope that a bit of his excellence has rubbed off on me over the years. The opportunity and privilege of being Rich's graduate student made the six long years in the cold and very Midwestern city of East Lansing well worth it.

Charles Ofria, my co-advisor, has been a great asset and help on the computational side of things, patiently introducing me to the world of Avida, Matlab, and C++. Also a wonderful friend, Charles never hesitated to offer a helping hand, whether it was a couch to crash on or a ride to the airport. Charles is an impressive rising-star scientist as well as a warm, open, and good-natured person. Just by being himself he has been an incredible role model, and will remain someone I look up to.

My committee members, Donald J. Hall and Robert T. Pennock readily jumped into the world of digital evolution bringing a diverse and substantial expertise. Their objective and novel views constantly challenged me, for which I am very thankful.

Chris Adami, my "academic grandfather", has left a significant imprint on my graduate career in spite of the geographical distance separating us. His probing and insightful comments and questions have helped me on many occasions and will continue to motivate me in the future.

Members of the Lenski and Ofria labs have created an uniquely challenging and supportive working environment, each and every person contributing to numerous discussions that have shaped my scientific thinking and growth. My thanks go to all of them: Elizabeth Ostrowski, Bob Woods, Tim Cooper, Kristina Hillesland, Chris Borland, Christopher Marx, Paco Moore, Susi Remold, Daniel Rozen, Gabriel Yedid, Sean Sleight, Zachary Blount, Brian Baer, Kaben Nanlohy, Neerja Hejela, Donna Ellis, Sherri Goings, Jeffery Clune, Jason Stredwick, Art Covert, Matthew Rupp, David Michael Bryson, Dehua Hang, Wei Huang, Larry Wisne, and Anastasha Kamps.

Never a true workaholic, I would have not been capable of completing my graduate studies without some quality down time. The diverse group of great friends that I have made in East Lansing kept me happy and sane, put up with my quirks, shared the good and bad times with me. While working on my laptop and drinking numerous espresso shots in the "Espresso Royal Cafe" (where these acknowledgments are being written) and relaxing at the neighboring "Peanut Barrel" (where I will go after these acknowledgements are completed, and do my best not to drink numerous non-espresso shots) I have found a community of incredible people that took me in and made my stay in East Lansing more fun than I expected or even hoped for. I could never list or truly thank them all and can just hope they remain in my life for many years to come: Dan Ducat and his family, Anthony Wittman, Ashley Nebe and her family, Elizabeth

Ostrowski, Bob and Andrea Woods, Keron Greene, Seth Walk, Chris Cobb and his family, Jason Bryans and his family, Billy Baldwin, Martha Sedivec, Billy Bewick, Tristan Zamp, Pete and Drew Weiss.

I have been gone from my home in Belgrade, Serbia for over 11 years and yet so many great friends still remain there: Sasa Novakovic, Aleksandar and Dubravka Antonijevic, Jelena Gaus, Jelena Milic, Ana Tomic, Ksenija Simic, Marjana Todorovic, Dejana Gajic, and Dusan Maljkovic. Always accepting me for who I am, they helped me though many hard times and are always in my heart and mind.

I must also express my thanks to the amazing educators and friends Borka Pesic, Vida Ognjanovic, Radmila Stefanovic, Tordis Isselhardt, and Jill Hayes. They thought me mathematics, literature, and English language but, more importantly, they recognized and nourished a potential in me that I hope I am beginning to fulfill.

My many relatives, most notably my aunts and uncles Petar and Lepa Hristovic and Dusanka and Dragan Nikolic, have sincerely celebrated and appreciated my academic achievements and have always lent a helping hand to my family and me.

My final words are about my parents Ljubica Hristovic Misevic and Mirko Misevic. They have not only unconditionally loved me and supported all my endeavors, but have also sacrificed so much for my benefit, always putting my future and well being first. My dad is now with me only in spirit but I know he would share my joy and happiness at this time. I remain forever indebted to my parents, humbled by the examples their lives set, and hopeful all my gratitude and love is at least somewhat reflected in my personal and professional accomplishments. I hope they are just half as proud to have me as their son as I am to have them as my parents.

TABLE OF CONTENTS

LIST OF TABLES ix

LIST OF FIGURES xi

CHAPTER 1
SEXUAL REPRODUCTION AND MULLER'S RATCHET IN DIGITAL ORGANISMS
- Abstract 1
- Introduction 2
- Materials and methods 4
  - The Avida system 4
  - First evolutionary stage 7
  - Second evolutionary stage 8
- Results and discussion 9
  - Adaptation and decay 9
  - Population survival and extinction 9
  - Mean fitness of surviving organisms 12
  - Distribution of mutational effects 14
  - Summary and future directions 15

CHAPTER 2
SEXUAL REPRODUCTION RESHAPES THE GENETIC ARCHITECTURE OF DIGITAL ORGANISMS
- Abstract 17
- Introduction 18
- Materials and methods 21
  - The Avida system 21
  - Modularity 23
  - Mutational sensitivity and epistasis 26
  - Statistics 27
- Results........ 27
  - Evolved sexual organisms have higher fitness and longer genomes 27
  - Evolved sexual organisms have more modular genomes 30
  - Mutational sensitivity and epistasis differ between sexual and asexual organisms........ 31
- Discussion 35
  - Relationships between the genetic architectural parameters 37
  - Implications for the evolution of sexual reproduction 39

CHAPTER 3
EXPERIMENTAL EVIDENCE OF EVOLUTION OF SEX IN CHANGING ENVIRONMENTS

Abstract 41
Introduction 41
Methods 43
Results and discussion 48

APPENDIX

Phenotypes in Avida 73
Resolving genome length issues 74
Reduced data set 75
Genome doublings prevented 76

REFERENCES 78

LIST OF TABLES

Table 1. Survival of sexual and asexual populations when population size is small and Muller's ratchet operates 11

Table 2. Distribution of one-step mutational effect on fitness in sexual and asexual ancestor of some second-stage populations 14

Table 3. Comparison of properties between sexual and asexual evolved populations 29

Table 4. Two-way ANOVA testing the effects of the period of environmental change and the substrate metabolic values on the proportion of time populations were dominated by sexual organisms 53

Table 5. Two-way ANOVA testing the effects of the initial mode of reproduction and period of environmental change on the proportion of time populations were dominated by sexual organisms 58

Table 6. Two-way ANOVA testing the effects of the initial mode of reproduction and period of environmental change on the proportion of time populations were dominated by sexual organisms 58

Table 7. Discriminant analysis for the final mode of reproduction under six periods of environmental change 61

Table 8. Discriminant analysis for the final mode of reproduction under six periods of environmental change 62

Table 9. One-way ANOVA testing the effects of the period of environmental change on the proportion of time populations were dominated by sexual organisms 69

Table 10. Linear and quadratic fits of proportion of time spent sexual to $\log_{10}$ transformed period of environmental change for populations evolving with 10 or 25 nutrients at a time 69

Table 11. Two-way ANOVA testing the effects of the mode of reproduction and the period of environmental change on mean fitness 71

Table 12. Truth table for nine logic operations 74

Table 13. Comparisons of properties between sexual and asexual evolved populations in the reduced data set 77

Table 14. Comparisons of properties between sexual and asexual evolved populations with genome doublings prevented 77

LIST OF FIGURES

Figure 1. Trajectories for maximum fitness and resource use during evolution in large and small asexual populations 10

Figure 2. Distribution of mean fitness in asexual and sexual populations that survived Muller's ratchet 13

Figure 3. Genotype-phenotype maps illustrating physical and functional modularity 25

Figure 4. Final distribution of average fitness (a) and genome length (b) in sexual and asexual populations 28

Figure 5. Genomic modularity of sexual and asexual organisms 33

Figure 6. Average fitness as a function of the number of one point mutations in evolved sexual (a) and asexual (b) organisms 34

Figure 7. Principal component analysis for average fitness, genome length, α, β, PM and FM of sexual and asexual digital organisms 38

Figure 8. Different periods of change in substrate metabolic values 47

Figure 9. Proportion of sexual organisms present in a population over time 50

Figure 10. Prevalence of sexual reproduction under 24 different treatments 52

Figure 11. Relationship between the prevalence of sex and the ancestral mode of reproduction 56

Figure 12. Proportion of time during experimental evolution that each population was predominantly sexual 57

Figure 13. Number of times each ancestral population evolved to be sexual 65

Figure 14. Prevalence of sexual reproduction in environments with different numbers of nutrients 67

Figure 15. Average fitness of predominantly sexual and predominantly asexual populations 70

CHAPTER 1

SEXUAL REPRODUCTION AND MULLER'S RATCHET IN DIGITAL ORGANISMS

ABSTRACT

The evolution of sexual reproduction has long been a major problem in biology. According to one theory, sex opposes the fitness-destroying process of Muller's ratchet, which occurs by the stochastic loss of high-fitness genotypes in small populations. Sex opposes the ratchet by allowing genotypes with different deleterious mutations to produce mutation-free offspring. We used the Avida digital-evolution software to investigate sex in relation to Muller's ratchet. Populations of digital organisms mutated, competed, and evolved in a complex environment. Populations were either asexual or sexual; in the latter case, parental genomes recombined to produce offspring. We also varied genomic mutation rates and population sizes, which at extreme values often caused mutational meltdowns and population extinctions. Our results demonstrate that sex is advantageous for population survival under some conditions. However, differences in extinction probabilities were usually small, occurred over a narrow range of mutation rates and population sizes, and the advantage of sex for population survival required many generations. Also, the mean fitness of surviving asexual populations was often greater than in sexual populations. This last result indicates the need for work that compares the statistical distribution of mutational effects and epistatic interactions in asexual and sexual populations.

INTRODUCTION

The mixing of genomes via sexual reproduction and the resulting recombination of genetic variants are widespread and important to life on Earth (Bell 1982; Maynard Smith 1978). For over a century, biologists have sought to explain the evolutionary origin and maintenance of sex (Ghiselin 1988; Weismann 1889; West et al. 1999). While there have been important advances, there are still more questions than answers, making this research area both interesting and active (Rice 2002).

There are many potential costs associated with sexual reproduction, including time and energy spent in searching for a suitable mate and in the act of mating (Lewis 1983). Moreover, at the genetic level there often exists a two-fold cost of sex, which is variously called the cost of meiosis or the cost of males. This cost alone implies that asexual populations should grow twice as fast as sexual ones, all else being equal (Maynard Smith 1971). This two-fold cost presents a major hurdle that must be overcome if the evolution of sex is to be understood. Many theories have been proposed to explain why sex arose and how it persists despite these costs, but most have not been adequately tested (Kondrashov 1993) and none are broadly supported by those tests that have been performed to date (West et al. 1999). One of the main theories is attributed to the famous geneticist H. J. Muller, and it proposes that sex is beneficial in opposing what is now referred to as Muller's ratchet.

Muller's ratchet is a stochastic process that leads to the loss of genetic information from asexual populations via the loss of those high-fitness genotypes that are free of any deleterious mutations (Felsenstein 1974; Muller 1964). The ratchet depends on both mutation and drift, where drift refers to changes in gene frequencies caused by the random sampling of a finite number of genes from the previous generation. Drift is especially important, and selection is relatively weak, in small populations that are most sensitive to sampling effects. By chance, the most fit genotypic class can be lost from a small population, especially if the genomic mutation rate is high. Once this class is lost from an asexual population, it cannot be recovered owing to the low probability of beneficial mutations in a small population. The metaphorical ratchet has thus advanced one notch and its action is irreversible; each subsequent loss of the most fit genotypic class advances the ratchet another notch. By contrast, in a sexual population the most fit class can potentially be reconstructed, even after it has been lost, via recombination between two genotypes that carry different mutations. The harmful effects of Muller's ratchet are thus opposed by sex.

Small populations subject to the ratchet may even go extinct if the resulting fitness decay causes the birth rate to drop below the death rate. This feedback can produce a vicious cycle, in which declining fitness leads to a drop in population size, which speeds up the ratchet causing further fitness loss. This cycle has been described as a mutational meltdown (Lynch et al. 1993; Lynch et al. 1995). In this study, we use population survival as one metric to compare the effect of Muller's ratchet in sexual and asexual populations.

Several experiments have shown the fitness-destroying effect of Muller's ratchet in diverse microorganisms including viruses (Chao 1990; Duarte et al. 1992), bacteria (Andersson and Hughes 1996), and protozoa (Bell 1989). At least one study with viruses further showed that genetic recombination could oppose the ratchet (Chao and Tran 1997). Research in evolutionary computation has also examined the ratchet, with an emphasis on designing strategies to eliminate its adverse effects from applied optimization algorithms (Laumanns et al. 2001; Nowak and Schuster 1989; Pürgel-Bennett 1997; Zitzler et al. 2000). However, questions about the operation of Muller's ratchet and its interaction with reproductive mode are difficult to answer using biological systems (Kondrashov 1982; Maynard Smith 1988). In this paper, we therefore use digital organisms to examine the effects of Muller's ratchet over a range of population sizes and mutation rates, and we test whether sexual reproduction can substantially impede the ratchet's harmful effects.

MATERIALS AND METHODS

The Avida system. We use digital evolution software called Avida to study Muller's ratchet and its interaction with reproductive mode. Avida maintains and monitors experimental populations of digital organisms, which are self-replicating computer programs written in a customized assembler-like language (Adami 1998; Ofria and Wilke 2004). Avida has been used for several other studies of evolutionary dynamics and outcomes (Lenski et al. 1999; Lenski et al. 2003; Wilke et al. 2001). Digital organisms in Avida evolve on rugged fitness landscapes characterized by diverse epistatic interactions, they have intricate genotype-phenotype maps that emerge from a complex developmental program, and they can exhibit quasi-species dynamics (Adami et al. 2000; Lenski et al. 1999; Lenski et al. 2003; Wilke et al. 2001). They approach the level of complexity of organic viruses, making results obtained with Avida biologically relevant and of general interest (Adami 2002; Wilke and Adami 2002). In Avida, genomes have 26 possible instructions at each position. All organisms descend from an ancestral program used to seed a population. Organisms execute the programs encoded by their genomes, including commands that enable them to copy and divide their genomes. The copy instruction duplicates a single instruction. During this duplication process, the instruction has a probability of being miscopied and changed to a different instruction in the offspring's genome; mutations from one instruction to any other are equally likely. In this study, we held the genome length constant by setting rates of insertion and deletion mutations to zero. The genomic mutation rate, U, equals the mutation probability per instruction copied multiplied by the genome length. The value of U is controlled by the investigator and was varied in our experiments.

Each digital organism occupies a cell in a rectangular lattice. The size of the lattice sets the maximum population size, which was also varied in our experiments. After a divide instruction is executed, the genome is split into two; the duplicated genome (the offspring), is placed into a random cell in the lattice, which kills the organism that previously occupied that position. Although death is random, the danger of being overwritten provides a selective advantage to organisms that replicate faster. Also, if any organism has not reproduced after executing its instructions an average of 15 times each, it dies and is removed from the population. If all the individuals in a population fail to reproduce within this allotted time, then the population has become extinct. In this study, organisms could accelerate the execution of their genomic instructions, and thus their reproduction, by performing certain logic functions (Lenski et al. 2003). If an organism performs one of these functions, then it receives some corresponding resource that provides energy and accelerates execution of its genomic program. Aside from differences in their ability to perform logic functions, all organisms would execute their genomes at the same rate. Even in that case, fitness can vary among organisms depending on their relative gestation time (number of executed instructions necessary to produce an offspring). An organism's expected fitness equals the product of the baseline energy available to all organisms (made proportional to genome length to eliminate selection on genome size per se) and bonuses received for performing logic operations, divided by the gestation time. Organisms do not have access to, and cannot manipulate, their expected fitness. Realized fitness is affected, however, by population structure and interactions among organisms.

In this study, we introduce a new command to Avida that causes the digital organisms to reproduce sexually. We use this variant command to compare evolution in asexual and sexual populations. When executed, the `divide-sex` command separates a copied genome from its parent, but it does not immediately place that new genome into the population. Instead, the new genome goes into a separate location called the *birth chamber*. If the chamber is empty, the new genome remains there until a second genome arrives. When two genomes are present, they recombine and then both resulting offspring are placed at random into the population. (Notice that this mechanism for sexual reproduction does not involve the two-fold cost of sex, although we could have introduced such a cost by placing only one of the two recombinant offspring in the population. It is likely that the most primitive biological forms of sex did not have to overcome this two-fold cost (Maynard Smith 1978), and so we began this research by placing both recombinants in the population. Even so, as we show below, the conditions favoring sexual reproduction with respect to Muller's ratchet are fairly narrow.) Recombination occurs by taking a single continuous region (with two random endpoints) from one genome and swapping it with the corresponding region from the other genome. Genomes are circular and fixed in length; genomic positions are defined by distance from the first command executed and direction of execution. The initial speed of execution of an offspring's genome is set to the weighted average of its two parents, with weights based on the proportion that each parent contributed to the offspring's genome. Under asexual reproduction, the initial speed is inherited from the sole parent.

We performed the evolution experiments with Avida in two stages. Briefly, the *first stage* used large populations in order to evolve digital organisms that were well adapted to their environment. For the *second stage*, these organisms were moved into much smaller populations to examine the effects of Muller's ratchet. Maximum population size was identical in all cases during the first stage, but this size was varied in the second stage. Mutation rates were also varied across runs, but the rate was held constant in both stages of a given lineage. Similarly, reproductive mode (asexual or sexual) varied across runs, but this mode was held constant during both stages of any lineage. Further details on the two evolutionary stages are provided below.

First evolutionary stage. All runs started with a hand-written ancestor, which had a genome of 100 instructions. The ancestor was capable of self-replication, but it could not perform any logic functions. Ten replicate experiments were run with each of five genomic mutation rates ($U = 0.1, 0.3, 1, 3, 10$) and with reproduction being either strictly asexual or sexual for the population. Thus, there were 100 runs of the first stage. Replicates differ only in the random number seed, which then affects all the stochastic events during the run, such as mutations and offspring placement. The maximum population size (N) was 3600 organisms for all runs in the first stage. Genome length was held constant in all runs. Nine different resources could be obtained by digital organisms that evolved the ability to perform logic functions; these resources were available in infinite supply. Experiments ran for 100,000 updates, where an update is an arbitrary unit of time in Avida corresponding to the execution of 30 instructions, on average, per individual organism. In these first-stage runs, one generation required roughly 10 updates; the exact value depends on the number of instructions needed to produce an offspring, which often changes during evolution. At $U = 0.1$ and 0.3, the experiments ran for an additional 500,000 updates in order to compensate for the slower adaptation at these lower mutation rates; this extension ensured there were genotypes that could use all nine resources in each first-stage treatment. During each run, we recorded the numbers of organisms using each resource as well as the mean and highest fitness in the population. At the end of each run, we saved the most fit genotype (provided it was able to use all nine resources) for use in the second evolutionary stage.

Second evolutionary stage. These runs used the pool of well-adapted genotypes from the first evolutionary stage as starting material to investigate the effect of Muller's ratchet on small populations. Each small population had the same mutation rate and same reproductive mode as its first-stage progenitor. For each of the five mutation rates, one sexual and one asexual organism were randomly chosen from the pool of genotypes saved at the end of the first-stage runs. Each of these ten genotypes (also referred to as proximate ancestors) was then used to start 100 replicate experiments at each of six small population sizes ($N = 4, 8, 16, 32, 64$, or 128), for a total of 6000 second-stage runs. All parameters other than population size were identical to those in the first-stage experiments. The second-stage runs lasted for 500,000 updates, during which we again recorded the number of organisms using each resource, as well as the mean and highest fitness. We also recorded whether the population had gone extinct by the end of the run.

RESULTS AND DISCUSSION

Adaptation and decay. In the first stage of the evolution experiments, several populations at each mutation rate produced highly adapted genotypes that could perform all nine logic functions and thereby obtain the corresponding resources. Figure 1a shows the maximum-fitness trajectory over the first 10,000 updates for a representative first-stage population that evolved with $U = 0.3$ while reproducing asexually. Figure 1b shows the corresponding trajectory for a second-stage population founded by the most fit genotype from the first stage, but with the maximum population size now reduced from 3,600 to 16 organisms. The step-like changes in maximum fitness are typical of the experiments. The steps reflect, in large measure, the adaptive gains and maladaptive losses of logic functions that occurred in the large and small populations, respectively (Figures 1c and 1d).

Population survival and extinction. In order to test if sexual reproduction could substantially impede Muller's ratchet, we compared the number of sexual and asexual populations that survived to the end of the second stage. Recall that mutation accumulation by Muller's ratchet can cause individuals to fail to reproduce and die, leading to a decline in population size which, if severe, might cause a mutational meltdown and eventual extinction. In total, we compared the fate of asexual and sexual populations under 30 different combinations of mutation rate and population size (Table 1). For 13 combinations with relatively large population sizes, low mutation rates, or

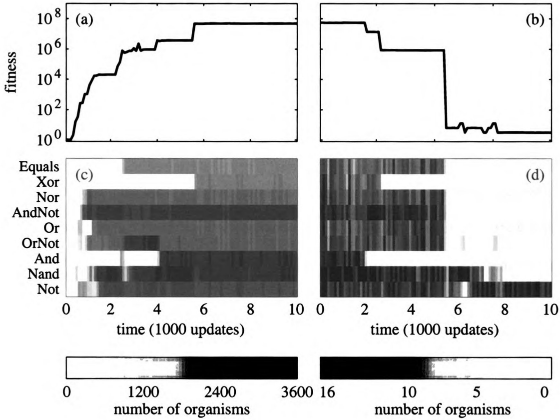


Figure 1. Trajectories for maximum fitness and resource use during evolution in large and small asexual populations. **(a)** Maximum fitness in a large population (N = 3600) during the initial 10,000 updates with genomic mutation rate U = 0.3; **(b)** Maximum fitness in a small population (N=16) that began with the most fit genotype from (a), and which continued at the same mutation rate; **(c)** Number of organisms performing each of the nine rewarded logic functions, indicated by shading intensity (scale below), in the same large population as in (a); **(d)** Number of organisms performing these logic functions in the same small population as in (b). Note the different scales in (c) and (d).

both, all 100 asexual and all 100 sexual populations survived to the end of the experiment. In 4 combinations subject to both high mutation rates and small population sizes, all 200 populations went extinct, regardless of their reproductive mode. In the remaining 13 combinations of mutation rate and population size, the number of surviving sexual populations was greater than the number of surviving asexual populations. For 4 of these combinations, the difference was significant based on Fisher's exact test (two-tailed $P < 0.05$) with a Bonferroni correction to adjust for performing 30 tests (Sokal and Rohlf 1995). Also, the likelihood that all 13 cases in which there was a difference would,

Reproduction: asexual, sexual		Genomic mutation rate, U				
		0.1	0.3	1	3	10
Population size, N	4	75, 80	**8, 41**	0, 1	0, 0	0, 0
	8	98, 100	92, 98	**60, 87**	0, 3	0, 0
	16	100, 100	100, 100	98, 99	**67, 95**	0, 0
	32	100, 100	100, 100	100, 100	97, 99	**18, 70**
	64	100, 100	100, 100	100, 100	99, 100	98, 99
	128	100, 100	100, 100	100, 100	100, 100	100, 100

Table 1. Survival of asexual and sexual populations when population size is small and Muller's ratchet operates. The two numbers in each cell show the number (out of 100 runs) of surviving asexual and sexual populations, respectively, for each combination of genomic mutation rate and maximum population size. Paired values are shown in **bold** when they were significantly different (see text for details).

by chance, trend in the same direction is very small (binomial test, $P < 0.001$). The data on population survival and extinction therefore support the hypothesis that sexual reproduction can slow the advance of Muller's ratchet and prevent mutational meltdown. On the other hand, most of the differences in extinction probabilities are fairly small, they depend on the particular parameter values for mutation rate and population size, and the survival advantage to sex requires thousands of generations to be manifest.

Mean fitness of surviving organisms. While the data on population survival are consistent with the hypothesis that sex is beneficial in opposing Muller's ratchet, the mean fitness values of survivors suggest a more complicated picture. Owing to the large number of experiments (100 populations for each of 60 combinations of population size, mutation rate, and reproductive mode), we cannot present all of the fitness data. However, Figure 2 shows the most important patterns. All of the populations in this figure evolved with genomic mutation rates set to 0.3; the three panels show data obtained for population sizes of 4, 16, and 64. Fitness values are expressed relative to the proximate ancestor, and were transformed owing to their tremendous range. At the lowest population size (Figure 2a), surviving sexual populations had slightly higher mean fitness values than did their surviving asexual counterparts, although this difference was not significant (two-tailed t-test, $P > 0.5$). For both reproductive modes, the final mean fitness values were very low relative to the ancestors. The situation was more complicated, however, at somewhat larger population sizes (Figures 2b and 2c). As expected, the mean fitness under both reproductive modes was much higher at $N = 16$ than at $N = 4$, and mean fitness was higher still with $N = 64$. Unexpectedly, however, the asexual populations had higher mean fitness than did the sexual populations at these

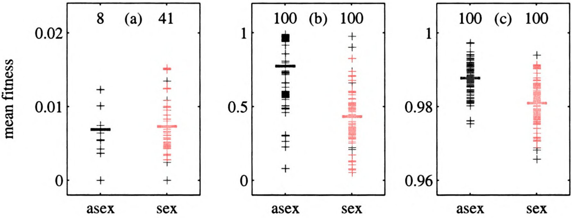


Figure 2. Distributions of mean fitness in asexual and sexual populations that survived Muller's ratchet. Panels **(a)**, **(b)**, and **(c)** correspond to maximum population sizes of 4, 16, and 64, respectively. All populations shown here evolved with a 0.3 genomic mutation rate. Asexual and sexual populations are shown at the left and right, respectively, in each panel. Each + symbol shows the mean fitness of one surviving population; the horizontal bar shows the mean value across the surviving populations for each treatment. The numbers of surviving populations in each treatment are shown along the top (see Table 1). Owing to the tremendous range of fitness values within and between treatments, mean fitness, W, is transformed as log(W+1)/log(W_0), where W_0 is the fitness of the proximate ancestor. Note the changes in fitness scale between panels. Images in this dissertation are presented in color.

larger population sizes (two-tailed t-test, both $P < 0.0001$). In appears that sexually reproducing populations, while better able to survive Muller's ratchet in very small populations, may accumulate more harmful mutations than asexual populations at somewhat larger population sizes. Data obtained from other combinations of population size and mutation rate gave similar results.

Distribution of mutational effects. As a first effort to understand one factor that might have contributed to lower mean fitness of sexual than asexual populations, we subjected the proximate ancestors to an in-depth mutational analysis. We constructed all 2500 one-step mutants (25 alterative instructions at each of 100 genomic sites) for the first-stage sexual and asexual genotypes that were ancestral to the populations shown in Figure 2. Table 2 shows the fraction of one-step mutations that were lethal, deleterious (non-lethal), neutral, and beneficial. The asexual genotype had a substantially higher proportion of mutations that were lethal, while the sexual type had a correspondingly

	Lethal	Deleterious	Neutral	Beneficial
Sexual	0.2420	0.6392	0.1088	0.0100
Asexual	0.4236	0.5052	0.0708	0.0004

Table 2. Distribution of one-step mutational effects on fitness in sexual and asexual ancestors of some second-stage populations. The proportion of the 2500 different mutations are shown for one sexual genotype and one asexual genotype that evolved during the first stage with $U = 0.3$, and which then served as the ancestors for the second-stage evolution shown in Figure 2.

higher proportion of deleterious but non-lethal mutations. Those populations derived from this asexual genotype would have had a higher risk of extinction, especially in the smallest populations, as a consequence of the higher fraction of lethal mutations. But surviving asexual populations might also have been purged of their most deleterious mutations, leaving these survivors with higher fitness than those from the sexual populations. The generality of these differences as a function of reproductive mode remains to be seen, as do such other factors as the extent and form of epistatic interactions between mutations (Lenski et al. 1999). But these preliminary data do suggest that prior evolution under the different reproductive modes can influence subsequent evolution. In other words, there are multiple interacting and dynamical feedbacks that shape evolving genomes, and they will complicate efforts to discern the various forces responsible for the origin and maintenance of sexual reproduction (Lenski 1999).

Summary and future directions. Sexual reproduction has several disadvantages relative to asexual reproduction, which begs the question of why sex is common in nature. Many potential advantages of sex have been hypothesized, including that sexual reproduction opposes the maladaptive effect of Muller's ratchet in small populations. We used the Avida software to perform evolution experiments with digital organisms that would test this hypothesis. Our results demonstrate the effect of Muller's ratchet in small populations. At high mutation rates and in very small populations, the ratchet often led to mutational meltdowns caused by the vicious cycle of mutation accumulation and population decline. Sexual populations survived this effect significantly better than asexual populations, but only over a fairly narrow range of parameter values. Opposing this advantage, surviving organisms in asexual populations unexpectedly had higher mean fitness than those in sexual populations at some other parameter values. This last result points toward the need for systematic analyses of the effect of reproductive mode on genetic architecture, including the distribution of mutational effects on fitness as well as the extent and form of epistatic interactions among mutations. Avida is well-suited for such analyses, which we intend to pursue in our future work on the evolution of sex.

CHAPTER 2

SEXUAL REPRODUCTION RESHAPES THE GENETIC ARCHITECTURE OF DIGITAL ORGANISMS

ABSTRACT

Modularity and epistasis, as well as other aspects of genetic architecture, have emerged as central themes in evolutionary biology. Theory suggests that modularity promotes evolvability, and that aggravating (synergistic) epistasis among deleterious mutations facilitates the evolution of sex. Here, by contrast, we investigate the evolution of different genetic architectures using digital organisms, which are computer programs that self-replicate, mutate, compete, and evolve. Specifically, we investigate how genetic architecture is shaped by reproductive mode. We allowed 200 populations of digital organisms to evolve for over 10,000 generations while reproducing either asexually or sexually. For ten randomly chosen organisms from each population, we constructed and analyzed all possible single mutants as well as one million mutants at each mutational distance from two to ten. The genomes of sexual organisms were more modular than asexual ones; sites encoding different functional traits had less overlap and sites encoding a particular trait were more tightly clustered. Net directional epistasis was alleviating (antagonistic) in both groups, although the overall strength of this epistasis was weaker in sexual than asexual organisms. Our results show that sexual reproduction profoundly influences the evolution of the genetic architecture.

INTRODUCTION

Modules are ubiquitous in biological systems, and they appear to be a critical aspect of biological organization (Hartwell et al. 1999; Schlosser and Wagner 2004). Defined as groups of characters serving the same function that are integrated into a unit largely independent from other such units (Wagner 1996), modules occur in such diverse contexts as the Hox gene cluster (Ferrier and Holland 2001), butterfly wing development (Beldade et al. 2002), and metabolic networks in bacteria and eukaryotes (Ravasz et al. 2002). Modular organization of genomes may facilitate the exchange of independent 'building blocks' via recombination, increase phenotypic variability, and promote evolvability (Kirschner and Gerhart 1998). While sexual reproduction with obligatory recombination is also common (Bell 1982), the selective forces responsible for its origin and maintenance – in the face of substantial costs – remain largely unknown even after a century of discussion and investigation (Kondrashov 1993; Maynard Smith 1978; Michod and Levin 1988; West et al. 1999; Williams 1975). One hypothesis emphasizes genetic architecture and suggests that aggravating (synergistic) epistasis between deleterious mutations may favor sex (Kondrashov 1982; Wolf et al. 2002), but various experiments show no excess of aggravating relative to alleviating epistasis (Chao 1988; De Visser et al. 1997; Elena and Lenski 1997; Lenski et al. 1999; Wilke et al. 2003). While these and other studies have considered the effects of genetic architecture on the evolution of sex (Kondrashov 1982; Rice and Chippindale 2001), the effects of reproductive mode on genetic architecture have received less attention (but see Lawrence and Roth 1996; Lenski 1999; Malmberg 1977; Pál and Hurst 2003; Pál and Hurst 2004). Here, by contrast, we measure both modularity and epistasis in the genomes of sexual and asexual evolving computer programs. We expect sex to promote more modular genomes, which may accelerate the origin of new traits and avoid disrupting existing traits. We expect sexual and asexual reproduction to yield different patterns of epistasis for two reasons. First, sex favors those mutations that enhance fitness across different genetic backgrounds, while asexuality favors mutations that are beneficial in the background in which they occur. Thus, stronger epistatic tendencies may evolve in asexual rather than in sexual reproduction (Malmberg 1977). Second, sex may promote aggravating epistasis relative to alleviating epistasis, because recombination would then facilitate the efficient removal of deleterious mutations. This explanation has been proposed for the evolution of sex (Kondrashov 1982), but the causal link might also be reversed.

In this study, we used ‘digital organisms’ to examine the effects of reproductive mode on the evolution of genetic architecture. These digital organisms are computer programs that replicate, mutate, and evolve in populations maintained by the Avida software (Ofria and Wilke 2004; Wilke and Adami 2002). They can perform various functions by executing the series of instructions encoded in their genomes, including instructions that enable them to copy their genomes line by line and thereby reproduce. Depending on the genetic program encoded by their genome, instructions may be executed out of order or multiple times. Point mutations, insertions, and deletions occur randomly during this process. Organisms compete for the energy they need to execute their genomic programs, and the resulting selection acts on heritable differences in their performance that are generated by mutations and, in sexual populations, recombination. Evolution therefore modifies the genome, with selection tending to reduce the number of instructions that must be executed to reproduce while increasing the energy available for execution. Organisms can augment the basal energy obtained at birth by performing computations in a manner analogous to metabolizing resources. In this study, nine distinct resources occur in unlimited quantities in the environment, but an individual can make use of each resource only once during its life cycle. The ancestor cannot perform any of these computations, but populations can evolve the ability to perform them.

Several studies of evolutionary dynamics and outcomes have taken advantage of the speed of evolution, flexible experimental design, and extensive data that can be obtained with Avida (Adami et al. 2000; Chow et al. 2004; Lenski et al. 1999; Lenski et al. 2003; Misevic et al. 2004; Wilke et al. 2001). Here, we examine the evolution of two key features of genetic architecture – modularity and epistasis – as a function of reproductive mode. To do so, we extended Avida to allow the possibility of sexual reproduction. Following the requisite site-by-site copying of a genome, the asexual divide instruction performs a genomic division and places the offspring into the population. The new sexual divide instruction requires the exchange of genetic material between two separately copied genomes before the recombinant offspring are placed in the population. Therefore, all offspring are the product of recombination under the sexual regime in Avida. We used two ancestors, capable of self-replication but not of performing any computations, and differing only in their divide instruction, to seed 100 sexual and 100 asexual populations in identical environments. Such experiments and subsequent analyses simply cannot be performed with any organic system at the scale, scope, and precision that digital organisms allow. We realize, of course, that the detailed results of similar experiments and analyses would undoubtedly differ between digital and organic systems (as they probably would also for different organic systems, if such work could be performed). Our intent, however, is to test general hypotheses about genome architecture in relation to mode of reproduction.

MATERIALS AND METHODS

The Avida system. We performed experiments with the Avida software, which is available without cost at http://devolab.cse.msu.edu/software/avida/. All experiments used default settings unless otherwise indicated. Reproductive mode is determined by whether the `divide-sex` or `divide-asex` instruction is included in the instruction set. All genomes used the default set of 25 instructions plus one of these divide instructions. Point, insertion, and deletion mutations occurred at rates of 0.002, 0.0005, and 0.0005 per instruction copied, respectively. Point mutations swap an instruction in the genome with a random one from the instruction set, while insertion and deletion mutations add or delete a random instruction. Besides their mode of reproduction, populations differed only in a random-number seed that affected all stochastic events during an experiment, including mutations and offspring placement. Each population had a maximum size of 3,600 organisms and was propagated for 100,000 updates. An update equals the time in which 30 instructions, on average, are executed per organism in the population. A typical generation is 5-10 updates, with the exact value depending on the complexity of organismal phenotypes, which changed during evolution. With asexual reproduction, an offspring is created by a division and immediately placed in the population. In sexual reproduction, the genome of an incipient offspring is first placed in a 'birth chamber' following a division. If the chamber is empty, the genome remains there until a second one arrives. (Sexual reproduction thus introduces a delay associated with the requirement for pairing. In principle, this delay might slow evolution in sexual populations relative to asexual ones. Pilot experiments in which some asexual populations experienced a corresponding delay, while others did not, showed that this delay had no appreciable effect on the rate of adaptation.) When two genomes are present, they recombine by swapping a continuous region of code. The relative locations of the beginning and end of that region are chosen from a uniform random distribution in the 0 to 1 range; the absolute locations are calculated by multiplying the random numbers by genome lengths. Genomes are circular and sites are numbered starting with the first one executed after birth. After recombination, both offspring are placed at random locations in the population. Organisms have no means of mate choice, and thus sexual selection is absent from this system. For both reproductive modes, offspring placement kills the organism that previously occupied that location, introducing genetic drift while maintaining a constant population size. All new organisms receive a basal amount of energy, which is scaled by genome length, thereby eliminating direct selection for small genomes. If, in the course of executing its genomic program, an organism inputs one or two 32-bit strings and outputs the result of one of nine basic logic operations performed in a bit-wise fashion on those strings, then it obtains additional energy at a rate roughly proportional to the operation's complexity. There are many different ways to perform each operation, and the number and identity of instructions used varies among organisms. The energy obtained by this computational metabolism is combined multiplicatively with the basal energy, and the product determines the relative speed with which each organism executes its genomic program. An organism's expected fitness is calculated as its total energy divided by the time used to produce an offspring and corresponds to the rate of offspring production. Fitness is expressed relative to the common ancestor; in our analyses of evolved populations, relative fitness was averaged over all organisms from a population and then $\log_{10}$ transformed.

Modularity. We obtained an organism's genotype-phenotype (GP) map by separately mutating each site in its genome and recording any resulting changes in the computational traits performed by the mutants. Physical deletions often produce confounding effects caused by changes in genome length; therefore, we mutated sites to an inert instruction (outside the default set available during evolution) that acts as a placeholder only. The GP map identifies those genomic regions that encode different traits. From the GP map, we then calculated two indices, physical and functional modularity.

Physical modularity, *PM*, measures the mean distance between sites encoding all computational traits:

$$PM = 1 - 2\frac{\sum_{t \in T} \sum_{i,j \in S_t, i \neq j} \frac{d(i,j)}{n(S_t)\cdot(n(S_t)-1)}}{L \cdot n(T)}$$

where L is genome length, T is the set of an organism's traits, $n(T)$ is the number of traits, S_t is the set of all sites encoding trait t, $n(S_t)$ is the number of sites encoding trait t, and $d(i,j)$ is the distance between sites i and j. The index averages the distance between two sites encoding a trait over the number of site pairs for each trait and the number of traits, and normalizes by genome length, which allows comparisons between organisms that differ in length and number of expressed traits. The normalized average is then doubled and subtracted from unity, yielding a score from 0 to 1, with higher values indicating more modular architectures. Characteristic distributions of instructions within a trait and the corresponding *PM* values are illustrated using hypothetical organisms in Figure 3a; evolved digital organisms generally require many more instructions to perform functional traits than these examples, which serve only to illustrate features of the *PM* metric. Organism 1 has the lowest possible modularity because the trait is encoded by only two instructions that are located as far apart from each other as possible given a circular genome. Organism 2 is more modular than 1 because, while the relevant instructions also lie in two distant regions, the multiple instructions within each region are very close together and thus lower the average distance. Organism 4 is more modular than 3 because all of the relevant instructions in 4 are within a single genomic region. Organism 5 has the highest physical modularity because the trait is encoded by just two adjacent instructions. While the *PM* for Organism 5 is not quite equal to the theoretical upper bound of 1, it would asymptotically approach that bound with increasing genome length.

Functional modularity, *FM*, measures the average overlap in the genomic sites that encode two or more different traits:

$$FM = \frac{\sum\limits_{a,b \in T, a \neq b} \sum\limits_{s \in S_a} (1 - m(s,b))}{L \cdot n(T) \cdot (n(T) - 1)}$$

where $m(s,b)$ specifies whether site s is required for expression of trait b, with $m(s,b) = \{1$ if $s \in S_b$; 0 if $s \notin S_b\}$. The number of non-overlapping sites for two traits is averaged over all trait pairs and normalized by genome length, again allowing comparisons between organisms that differ in length and number of expressed traits. Characteristic types of overlap between traits and the corresponding *FM* values are shown for hypothetical

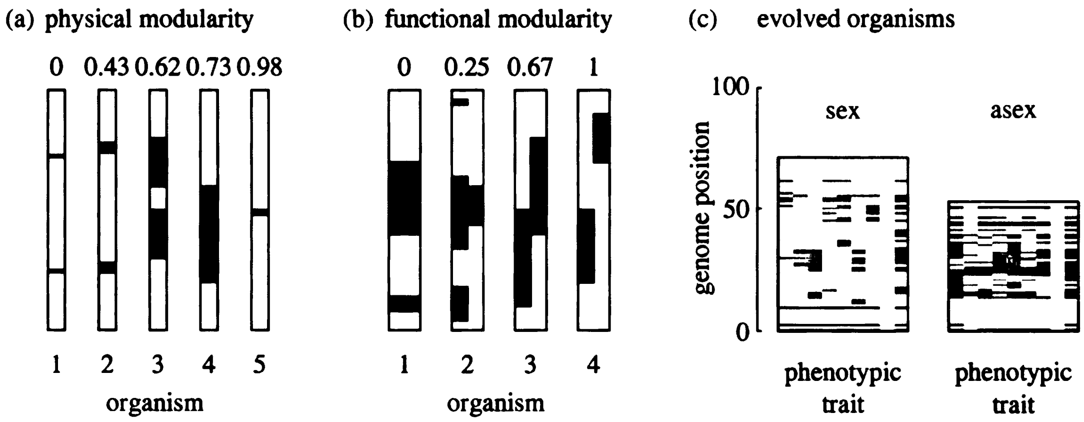


Figure 3. Genotype-phenotype maps illustrating physical and functional modularity. Phenotypic traits are arrayed as columns and genomic sites as rows. For simplicity, the circular genomes are shown in a linear fashion. An open cell indicates that the instruction at that site can be deleted without eliminating the trait; a filled cell marks the site of an instruction required for expression of the trait. If a mutation not only disrupts a trait but also prevents the organism from replicating, then that site is not filled in the GP map. **(a)** Five hypothetical organisms of equal length that express a single functional trait, with physical modularity values indicated above. **(b)** Four hypothetical organisms of equal length that express two traits each, with functional modularity values indicated above. **(c)** Representative GP maps for sexual (left) and asexual (right) organisms. *PM* and *FM* values for this sexual organism are 0.796 and 0.841, respectively; *PM* and *FM* values for the asexual organism are 0.689 and 0.746, respectively. Each value is within 5% of the overall mean for the corresponding reproductive mode.

organisms in Figure 3b. When there is complete overlap between the instructions encoding different traits, then *FM* is 0, as in Organism 1. By contrast, *FM* equals 1 when the instructions encoding different traits have no overlap at all, as in Organism 4. Intermediate values of *FM* are represented by Organisms 2 and 3. Organisms that do not express any computational traits (including those sampled before any functions evolved and, later, those with severe mutations) have undefined physical modularity and were excluded from the analysis. Similarly, *FM* is defined only for organisms expressing at least two functions because it measures the overlap in the instructions that encode them.

Modularity has joined the ranks of biological properties, such as fitness and species, that are often difficult to define and measure, resulting in a multitude of different and sometimes even conflicting definitions (Beldade et al. 2002; Ferrier and Holland 2001; Hartwell et al. 1999; Ravasz et al. 2002; Schlosser and Wagner 2004; Wagner 1996; Winther 2001). In pilot experiments, we explored several different modularity metrics ranging from quite simple (e.g., average number of tasks affected by an instruction) to more complex ones (e.g., *FM* further normalized by the expected overlap given the density of instructions in the GP map). Results obtained with all the metrics led to qualitatively similar conclusions. Therefore, we chose to focus on *PM* and *FM*, which measure different, yet intuitive and meaningful, aspects of genomic architecture.

Mutational sensitivity and epistasis. We mutated each site in an organism's genome to every alternative state (e.g., 50 × (26-1) = 1,250 mutants for length 50 genomes and 26 different possible instructions) and measured the resulting fitness. For each organism, we also examined random samples of a million mutants for each mutational distance from two to ten; these samples included more than 18-billion mutants in all. Parameters α and β define an organism's average sensitivity to individual mutations and the overall form of epistatic interactions among mutations, respectively (Elena and Lenski 1997; Lenski et al. 1999). We calculated each organism's α exactly as $-\log_{10}W(1)$, where $W(1)$ is the average fitness of all single mutants expressed relative to the unmutated state. We then estimated each organism's β by minimizing the sum of squared deviations between the actual average fitness values for $1 \leq m \leq 10$ and those predicted by the power function $\log_{10} W(m) = -\alpha m^{\beta}$, using α as obtained above.

Statistics. We used Systat 10.2 software (SSI, Richmond, CA) for all statistical tests.

RESULTS

Evolved sexual organisms have higher fitness and longer genomes. Sexual populations adapted significantly better to the environment, on average, than did asexual populations, as indicated by average fitness values after 100,000 updates ($P = 0.006$; Table 3; Figure 4a). Sexual populations also evolved larger genomes, on average, than asexual populations ($P < 0.001$; Table 3; Figure 4b), but this difference was strongly influenced by the fact that 35 sexual populations evolved genomes more than twice the ancestral length of 50 whereas only 9 asexual populations did so. Similarly, there was more diversity in genome length within individual sexual populations (data not shown), probably caused by the larger changes in genomes allowed by recombination. The large genomes typically evolved via genome doublings that occurred when, owing to some mutation, the genetic program failed to detect that the genome had already been copied

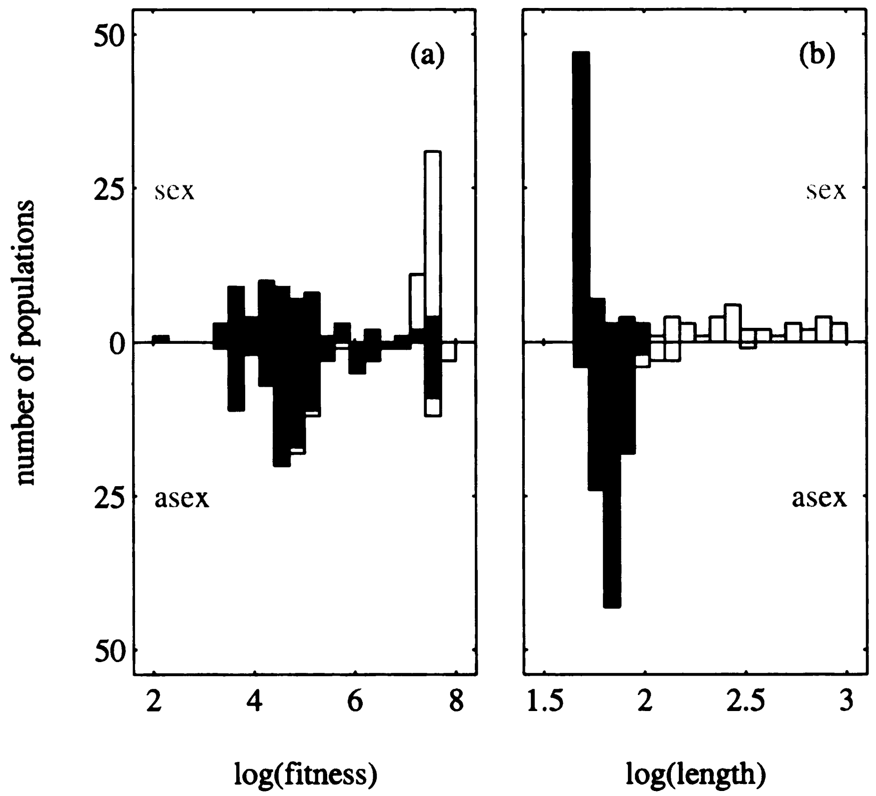


Figure 4. Final distributions of average fitness (**a**) and genome length (**b**) in sexual and asexual populations. Open sections show populations with average genome length ≥100 instructions.

and repeated the site-by-site replication. Sexual reproduction in Avida evidently increased the chance of genome doubling, created more favorable genetic combinations in the larger genomes, or both. To preclude genome doublings from possibly biasing our analyses of genetic architecture, we reduced the data set by excluding all populations with mean length ≥100 (i.e., at least twice the ancestral length; Appendix). We also performed additional experiments in which we actively prevented genome doublings from occurring (Appendix). In both cases, the resulting asexual populations had, on average, longer genomes than sexual ones, thus reversing the direction of potential bias. The alternative analyses also eliminated the significant fitness difference between sexual and asexual populations, with higher fitness shifting, albeit insignificantly, to the asexual

response variable	mean sexual (± s.d.)	mean asexual (± s.d.)	P
$\log_{10}$ fitness	5.762 (1.610)	5.198 (1.228)	0.006
$\log_{10}$ genome length	2.022 (0.414)	1.853 (0.114)	< 0.001
physical modularity, *PM*	0.824 (0.098)	0.699 (0.073)	< 0.001
functional modularity, *FM*	0.838 (0.091)	0.766 (0.079)	< 0.001
average effect of single point mutations on fitness, α	0.546 (0.292)	0.768 (0.190)	< 0.001
net directional epistasis, β	0.929 (0.057)	0.862 (0.088)	< 0.001

Table 3. Comparisons of properties between sexual and asexual evolved populations. The P values are based on two-tailed *t*-tests.

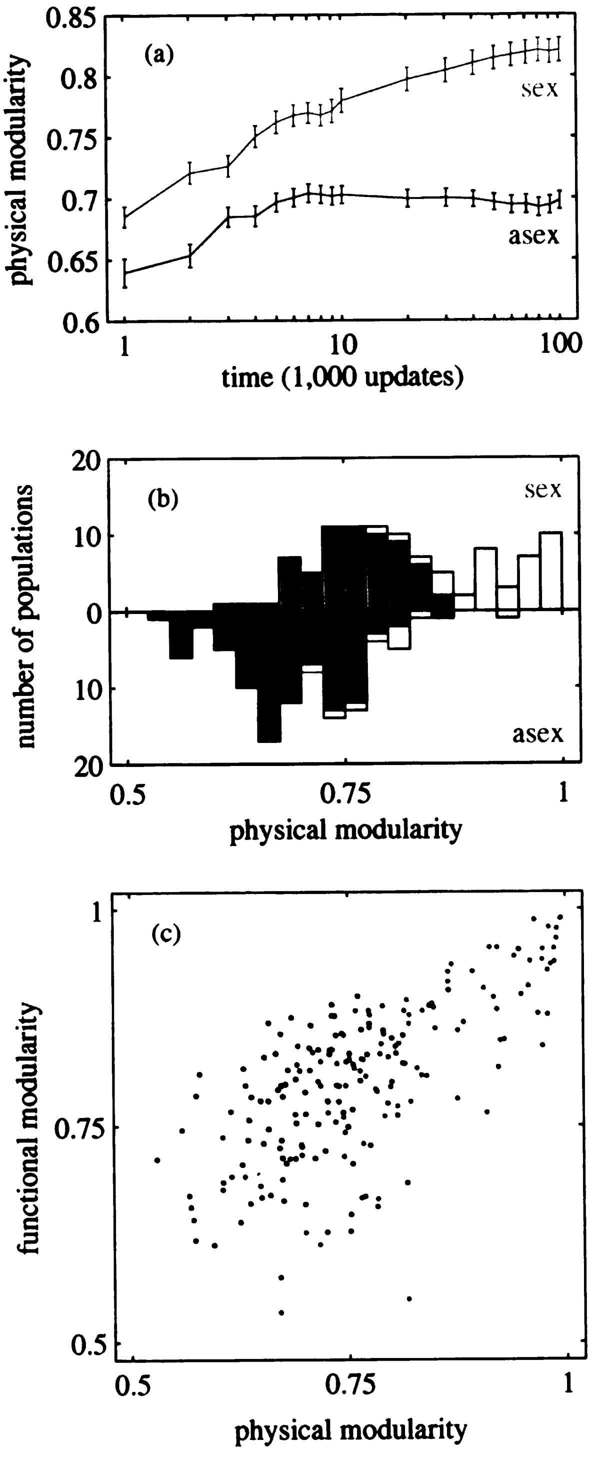
populations. Importantly, all of the qualitative effects of reproductive mode on genetic architecture remained in the same direction, and significant in all but one case, under both alternative analyses. Therefore, these architectural differences are robust with respect to differences in genome length between sexual and asexual organisms.

Evolved sexual organisms have more modular genomes. We performed all of the analyses of genetic architecture on a random sample of ten viable organisms from each population at the end of the evolution run; occasional non-viable genotypes were excluded. For each organism, we first deleted each genomic site (replacing it with an inert placeholder instruction) in order to construct genotype-phenotype (GP) maps; these maps identify the set of sites required to express a particular trait (compute a logic operation). We used the independently evolved populations as the unit of replication in statistical comparisons, because organisms sampled from the same population inevitably share much of their ancestry.

Using the GP maps, we then calculated physical modularity (*PM*), which reflects the average distance between two sites encoding a particular trait, and functional modularity (*FM*), which captures the average overlap between sites encoding different traits. *PM* and *FM* can range from 0 to 1, with high *PM* values indicating that traits are encoded in compact regions of the genome and high *FM* values corresponding to low overlap between traits. Representative GP maps illustrate the difference in modularity between sexual and asexual organisms: sites encoding the various traits in sexual organisms tend to be more compact (shorter, more continuous vertical filled blocks) and less overlapping (shorter, fewer horizontal filled lines) than those in asexual organisms (Figure 3). We averaged *PM* and *FM* across organisms within a population, and then compared the populations with different reproductive modes. Over time, the genomes of sexual organisms became increasingly more physically modular than asexual ones (Figure 5a), and this difference was highly significant at the end of the experiment ($P < 0.001$; Table 3; Figure 5b). The difference in *PM* remained highly significant when populations with genome length ≥ 100 were excluded from the analysis ($P < 0.001$; Appendix). Similarly, sexual organisms evolved genomes with a significantly higher *FM* than asexual ones ($P < 0.001$; Table 3, Figure 5c), a difference that also holds in the reduced data set ($P = 0.002$; Appendix).

Mutational sensitivity and epistasis differ between sexual and asexual organisms. We further examined genetic architecture by quantifying the form and extent of epistasis in the same organisms used to assess modularity. For each organism, we made all possible one-step point mutants and obtained random samples of a million organisms carrying from two to ten mutations; for each mutant, we measured its fitness relative to its unmutated parent. We averaged relative fitness over the organisms in a mutational class from each population. We then used a power function, $\log_{10} W(m) = -\alpha m^{\beta}$, to relate average mutant fitness, W, to the number of mutations, m. Here, α expresses the rate of change in average fitness expected if mutations acted independently, and β describes the overall form of epistasis. If $\beta = 1$, then mutational effects are on average multiplicative (no epistasis); if $\beta < 1$, then additional mutations tend to reduce fitness less than expected from their individual effects (alleviating epistasis); and if $\beta > 1$, then additional mutations reduce fitness more than expected from their individual effects (aggravating epistasis).

Figure 5. Genomic modularity of sexual and asexual organisms. (**a**) Trajectories of average physical modularity in evolving sexual (red) and asexual (blue) populations. Note the logarithmic time scale and that the first time-point represents 1,000 updates. Error bars are standard errors of the mean. (**b**) Final distributions of physical modularity values. Open sections show populations with average genome length ≥100 instructions. (**c**) Relationship between the final physical and functional modularity values. Each point represents a single independently evolved sexual (red) or asexual (blue) population.


(a)
physical modularity
0.85
0.8
0.75
0.7
0.65
0.6
sex
asex
1
10
100
time (1,000 updates)
(b)
number of populations
20
10
0
10
20
sex
asex
0.5
0.75
1
physical modularity
(c)
functional modularity
1
0.75
0.5
0.5
0.75
1
physical modularity

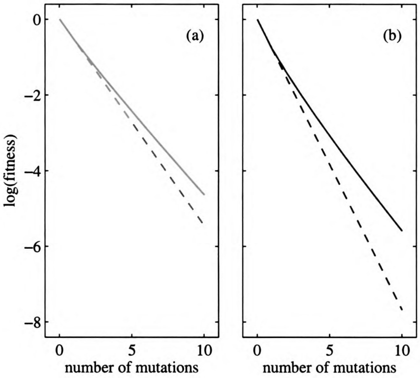


Figure 6. Average fitness as a function of the number of random point mutations in evolved sexual (**a**) and asexual (**b**) organisms. Dashed lines indicate the fitness decay functions expected under a multiplicative model of mutation interactions, using the corresponding average value for α and setting β = 1. Solid curves are the decay functions based on average values for both α and β parameters.

Sexual populations became more robust, on average, to individual mutations than did asexual ones (Figure 6), with sexual organisms having significantly lower α values ($P < 0.001$; Table 3). The predominant form of epistasis was alleviating in both sexual and asexual organisms ($\beta < 1$ based on t-tests, both $P < 0.001$), although this directional epistasis was weaker in sexual organisms ($P < 0.001$; Table 3). The differences between sexual and asexual populations remained significant even when those with genome length ≥ 100 were excluded ($P = 0.019$ for α, $P < 0.001$ for β; Appendix).

DISCUSSION

In nature, organisms show considerable variation in their reproductive mode as well as in certain features of their genetic architecture. While there has been substantial interest in the relationship between reproductive mode and features of genetic architecture, it is difficult to establish causation (Beldade et al. 2002; Chao 1988; De Visser et al. 1997; Elena and Lenski 1997; Ferrier and Holland 2001; Hartwell et al. 1999; Kirschner and Gerhart 1998; Lawrence and Roth 1996; Lenski 1999; Lenski et al. 1999; Malmberg 1977; Pál and Hurst 2003; Pál and Hurst 2004; Ravasz et al. 2002; Rice and Chippindale 2001; Schlosser and Wagner 2004; Wagner 1996; Wagner and Altenberg 1996; Wilke et al. 2003; Wolf et al. 2002). This difficulty reflects several challenges including a multitude of potential feedbacks between the biological features of interest, the inability to manipulate such critical features in most biological systems, and the infeasibility of performing long-term experiments to observe how manipulating one factor would impact the evolution of other features.

In this study, we have overcome these challenges by investigating the evolutionary relationship between reproductive mode and genetic architecture in digital organisms, which are self-replicating computer programs that mutate, compete, and evolve. By comparing evolutionary outcomes in populations that evolved in identical environments and initially differed only in whether they reproduced asexually or sexually, we demonstrated that several features of genetic architecture were shaped by reproductive mode. The genomes of sexual organisms were significantly more modular than those of asexual organisms by two different measures (Figure 5). Sexual organisms were also more robust with respect to the average effect of single mutations, while asexual organisms tended to have stronger epistatic interactions among multiple random mutations (Figure 6).

An unexpected complication arose because sexual organisms often evolved much longer genomes than asexual organisms (Figure 4b), evidently reflecting a greater propensity of the sexual populations to generate or retain genome doublings. Genome length was itself correlated with other features of genetic architecture such as modularity (Figure 5b), which raised the question of whether differences between sexual and asexual populations in genome length might be solely responsible for the other differences in their genetic architecture. We examined this issue in two ways (Appendix). First, we excluded from our analyses all evolved populations in which genome length had at least doubled. Second, we performed an additional set of experiments that prevented genome doublings and other large changes in genome size from occurring. In both cases, the difference in genome length became shifted in the opposite direction such that asexual organisms were in fact longer. Yet, the other differences in genetic architecture between sexual and asexual organisms retained their original form, with the genomes of sexual organisms being more modular and more robust to the fitness effects of individual mutations, while exhibiting weaker epistatic interactions among mutations.

Relationships among the genetic architectural parameters. Previous research has demonstrated negative correlations between mutational parameters α and β using both analytical models and simulations of RNA folding (Wagner et al. 1998; Wilke and Adami 2001). We, too, observe a strong negative correlation between α and β in the combined data set that includes both sexual and asexual digital organisms ($r = -0.694$, $P < 0.001$ for all 100 evolved populations, and $r = -0.613$, $P < 0.001$ for the reduced data set that excludes populations with genome length ≥ 100). We also examined the relationship between the two measures of modularity, *FM* and *PM*, and found that they are positively correlated ($r = 0.685$, $P < 0.001$, and $r = 0.491$, $P < 0.001$ for the reduced data set). All else equal, genomes with more compact regions expressing different traits would have less overlap if those regions were randomly distributed throughout a genome, which may contribute to this association. However, the observed correlation coefficient is only moderate in magnitude, and therefore we conclude that these two modularity measures capture somewhat different aspects of the genetic architecture.

To examine the relationships among all the genetic architectural and performance metrics in this study, we next performed a principal component analysis (PCA) on log-transformed fitness, log-transformed genome length, *PM*, *FM*, α and β (Figure 7a). The first two components explain over 70% of the total variance in the data. The first principal component, pc_1, largely reflects the two modularity measures, *PM* and *FM*, whereas pc_2 reflects the directional epistasis parameter, β, and fitness (Figure 7b). Both

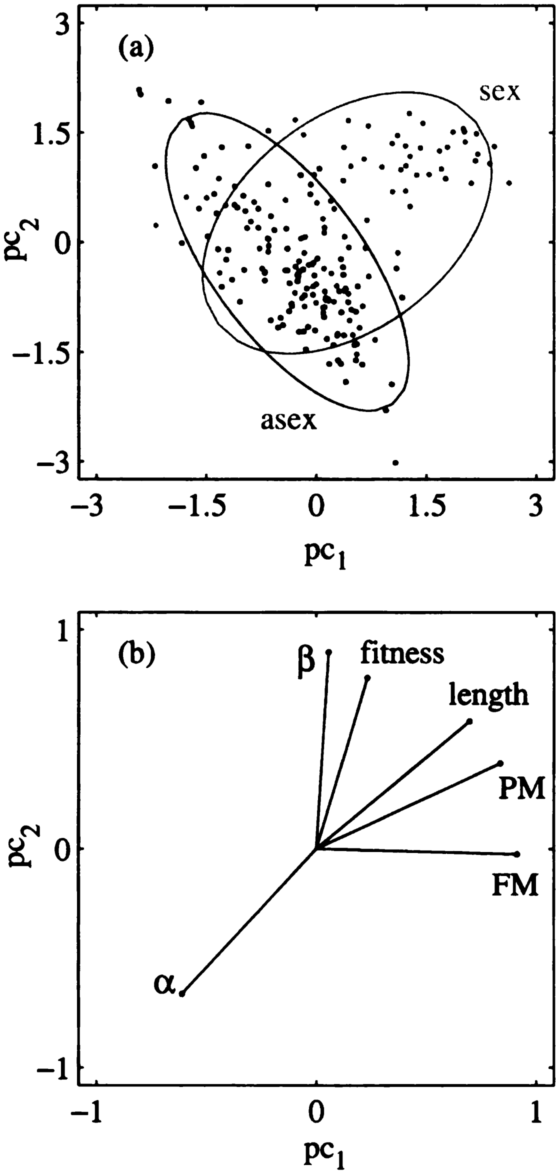


Figure 7. Principal components analysis for average fitness, genome length, α, β, *PM*, and *FM* of sexual and asexual digital organisms. (**a**) Transformed data for sexual (red) and asexual (blue) populations, with 95% confidence ellipses. (**b**) Loadings for the two principal components.

components also show substantial loadings for genome length and mutational sensitivity, α, with these two metrics being negatively correlated, such that longer genomes tend to be less sensitive, on average, to single mutations.

We also performed a discriminant function analysis using the same six properties as in the PCA in order to examine which ones best capture the observed differences between sexual and asexual organisms. The discriminant function gives the largest and nearly equal weights to *PM* and β, and it would correctly classify a given organism as sexual or asexual in over 80% of all cases (data not shown). We therefore conclude that epistasis and modularity evolve differently in sexual and asexual populations and, moreover, with sufficient independence that no single feature of the genetic architecture can explain all the differences between sexual and asexual organisms.

Implications for the evolution of sexual reproduction. Our experiments do not specifically address the long-standing question of why sex evolved (Bell 1982; De Visser et al. 1997; Elena and Lenski 1997; Kondrashov 1982; Kondrashov 1993; Lenski et al. 1999; Maynard Smith 1978; Michod and Levin 1988; West et al. 1999). Nevertheless, some of our findings bear on this issue. In particular, our results support previous studies [Chao, 1988; Wilke et al. 2003; Williams 1975) that failed to find a preponderance of aggravating (synergistic) epistasis, which is an essential component of the mutational deterministic hypothesis (Kondrashov 1982). Importantly, we extend this conclusion by showing that an excess of alleviating epistasis exists even in sexual populations (Figure 6). Our experiments also suggest two alternative hypotheses for the evolution of sex, one that is immediate and the other longer term in its consequences. First, sexual populations evolved lower values of α (Table 3), indicating reduced sensitivity to the usually harmful effects of individual mutations. Thus, sexual progeny are more robust than asexual progeny owing to differences in the genetic encoding of their phenotype (e.g., canalization), as opposed to the role of recombination in purging deleterious mutations that is posited under various other hypotheses (Kondrashov 1982). In this respect, it is important to recognize that distributions of mutational effects and forms of epistatic interactions evolved freely in our experiments as populations moved across fitness landscapes, whereas these quantities are usually fixed in population-genetic analyses. This finding suggests that theoreticians should give more attention to understanding the structure of fitness landscapes and how different modes of reproduction influence where populations settle on those landscapes. Second, an indirect benefit to sex could arise from higher genomic modularity, which may increase their evolvability and thereby promote greater fitness (Earl and Deem 2004; Wagner and Altenberg 1996). This hypothesis is more relevant to the maintenance of sexual reproduction than to its origin, however, because the benefit would not be manifest immediately (Bell 1982; Kondrashov 1982; Kondrashov 1993; Maynard Smith 1978; Michod and Levin 1988; West et al. 1999; Williams 1975). Finally, our results demonstrate that reproductive mode substantially shapes the evolution of the genetic architecture. We therefore emphasize the importance of distinguishing between the evolutionary causes and consequences of sexual reproduction.

CHAPTER THREE

EXPERIMENTAL EVIDENCE OF EVOLUTION OF SEX IN CHANGING ENVIRONMENTS

ABSTRACT

The evolution of sex is a vibrant and fascinating area of research, rich with theories but poor in experimental results. The crux of the sexual paradox is the dominance of the sexual mode of reproduction in the natural world in spite of the apparent costs associated with recombination. Here we demonstrate that sex can evolve *de novo* and outcompete the asexual mode of reproduction under changing environmental conditions. We evolved large populations of digital organisms for thousands of generations in six environments with different periods of substrate change. Sex was the dominant mode of reproduction when the environments were changing rapidly, with up to 65% of populations evolving to reproduce sexually. The ancestral reproductive mode and genetic architecture had only weak effects on whether populations evolved to be sexual or asexual while adapting to the rapidly changing environment. We also found that in the environmental conditions where sex was prevalent, the sexual populations on average had higher fitness than asexual ones. Our study experimentally demonstrates the importance of changing environments for the evolution of sex.

INTRODUCTION

Why sex? This complicated and often costly mechanism of reproduction dominates the biological world and continues to motivate a multitude of research studies, theoretical models, and diverse hypotheses (Bell 1982; Michod and Levin 1988; Rice 2002; Weismann 1889). The underlying benefit of sex has most often been related to the genetic variation that is created by recombination and to the faster adaptation resulting from that genetic variation (Maynard Smith 1978; West et al. 1999; Williams 1975). Faster adaptation is especially important when populations are facing a novel or changing environment. A number of experiments demonstrate the benefit of sex for adaptation to new or harsh environments. For example, sex increases the efficiency of selection during adaptation to a novel environment in yeast (Goddard et al. 2005), and a high recombination rate accelerates the evolution of drug resistance in bacteriophage T4 (Malmberg 1977), while a low recombination rate slows down the response to artificial selection in fruitflies (McPhee and Robertson 1970). Conceptually, evolution in a changing environment is similar to repeated bouts of adaptation to a novel environment. Much theoretical work indicates that changing environments promote the evolution of sex, but experimental support is sparse (Barton 1995; Charlesworth 1993; Kondrashov and Yampolsky 1996; Otto and Michalakis 1998; Sasaki and Iwasa 1987; Waxman and Peck 1999; Wolf et al. 1987). In an exceptional example from nature, host-parasite interactions create a biotically changing environment that modulates the frequencies of sexual and asexual snails and supports the popular Red Queen hypothesis (Lively 1987; Lively and Dybdahl 2000). Here, we perform experiments with a digital system and demonstrate that sexual organisms can invade and outcompete asexual ones when the environmental conditions are changing rapidly, thereby providing direct experimental evidence for a benefit of sexual reproduction.

METHODS

To test whether environmental change promotes the evolution of sex, we use the Avida experimental evolution software. Avida populations comprising of thousands of self-replicating computer programs are propagated inside computers. These digital organisms mate and reproduce, mutate, compete, metabolize some substrates while being poisoned by others, adapt, and evolve for thousands of generations. Avida captures evolution in a computer and has been used to study a variety of questions about evolutionary processes and outcomes. Most notably, previous studies have investigated the evolution of complex features, genetic architecture and speciation in digital organisms (Chow et al. 2004; Lenski et al. 2003; Misevic et al. 2006).

Avidians reproduce by executing the instructions present in their digital genomes. However, the organisms' genetic programs may also contain instructions that input, output, and manipulate random numbers. The environment contains multiple substrates that organisms can metabolize to gain additional energy and speed up their replication. Each substrate is associated with a particular mathematical logic operation. An organism receives an energy bonus if it evolves and executes a sequence of instructions that inputs random numbers, performs substrate-associated logic operation on those numbers, and outputs the result. Energy received via digital metabolism is combined with the organism's basal energy to determine its relative replication rate. We measure replication time and metabolic energy gains to determine fitness. As discussed below, certain substrates may also be toxic at some points in time and their metabolism would then decrease organism's energy level. Fitness values are averaged over all organisms in the population and $\log_{10}$ transformed.

Reproduction in Avida is typically asexual – organisms replicate their genomes and create offspring by simple division. Recently, recombination was introduced in the system (Misevic et al. 2004) and this is the first study in which sexual and asexual digital organisms mutate from one reproductive mode to the other and thus directly compete in the same population. During sexual reproduction, two sexually generated offspring randomly pair up and exchange a continuous region of their genomes. The physical position of the exchanged region within the genomes is equivalent and its size is proportional to the parental genome length. After the recombination process is complete, both incipient offspring are placed at random locations in the population. A single instruction in the organisms' genetic program performs the division and determines the mode of reproduction. The offspring produced by the `divide-sex` instruction will undergo recombination, while the ones produced by `divide-asex` will bypass the recombination step and immediately start their lifecycle. We can limit populations to being purely sexual or asexual by including only one of the two possible divide instructions in the mutationally accessible instruction set. However, if mutations to and from both divide instruction are possible, as they were in all the experiments here, we can directly examine evolution and competition between organisms with different modes of reproduction.

In Chapter 2 (Misevic et al. 2006) fifty ancestral populations evolved for over 10,000 generations under obligatory sexual or obligatory asexual conditions while their fitness, genome length, and genetic architecture were recorded. Here, we continued propagating those populations and preserved all but two experimental settings. Firstly, digital organisms in any population can evolve to be either sexual or asexual through mutations. Secondly, the environments in which populations evolve are continuously changing. The environments are characterized by the substrates they contain, and the environmental change is caused by the fluctuation in the metabolic values of those substrates. The range of metabolic values a substrate can have and the time between changes in a substrate's metabolic value, are the two aspects of the environmental change that we consider here. The environment in Chapter 2 contained 9 substrates, with metabolic values 1, 1, 2, 2, 3, 3, 4, 4 and 5, with higher values associated with more complex computational functions. When an organism metabolizes a substrate with a value 3, for example, it receives 2^3-fold increase in its current energy level. The organism's energy level determines the speed of execution of its genetic program and thus its speed of replication. All substrates in Chapter 2 had positive metabolic values, making their metabolism beneficial and so we designate them as nutrients. However, substrates may also have negative metabolic values, and their consumption would then slow down the organism's reproduction, making those substrates poisonous to the organism. In this study, all environments contained the 9 substrates present in Chapter 2; those substrates were always nutrients, with the same metabolic values as in Chapter 2. In addition, 68 more substrates were present in environments in this study. Those substrates took one of two possible metabolic values, "stick" (s) and "carrot" (c), with $s < 0$ and $c > 0$. During evolution, 25 randomly chosen substrates were always nutritious, while the remaining 43 were poisonous; however, the identity of those substrates changed randomly over time. All 68 changing substrates in a particular environment share the same (s, c) values. We evolved populations in environments with four different (s, c) values: (-1, 1), (-1, 3), (-1, 5), and (-3, 1). The substrates that are sometimes nutritious and sometime poisonous provide a two-fold impetus for adaptation: (1) energy gained from metabolizing a nutritious substrate ("carrot") and (2) energy lost from metabolizing a poisonous one ("stick") which may have been nutritious a short while ago.

Besides the metabolic values of the substrates, each changing environment is characterized by the period of time between changes in those values. The environments change gradually, by a swapping the metabolic values of two randomly chosen substrates at a time. For example, at the environmental change period of 10u, every 10 updates one of the 25 nutritious substrates becomes poisonous and one of the 43 poisonous substrates becomes nutritious. The change from a nutrient to a poison is achieved simply by changing the substrate's current metabolic value from c to s. An update equals the time in which, on average, 30 genomic instructions are executed per organism in the population. A typical generation is 5-10 updates, with the exact value depending on the complexity of organismal genotypes and phenotypes, which may change during evolution. Here, we studied evolution in environments with six different periods of change: 1u, 3u, 10u, 30u, 100u, and 300u (Figure 8). For each of the four metabolic value pairs and each of the six periods of environmental change, we propagated 20 initially sexual and 20 initially asexual populations from Chapter 2 for additional 100,000 updates, a total of 4 x 6 x 2 x 20 = 960 populations. Over the course of evolution, many of those populations switched their mode of reproduction, which we recorded.

All the experiments were performed using Avida software, which is available without cost at http://devolab.cse.msu.edu/software/avida/. We used the default settings, unless otherwise indicated. The organism's mode of reproduction was determined by whether it executed `divide-sex` or `divide-asex` to reproduce. All genomes used

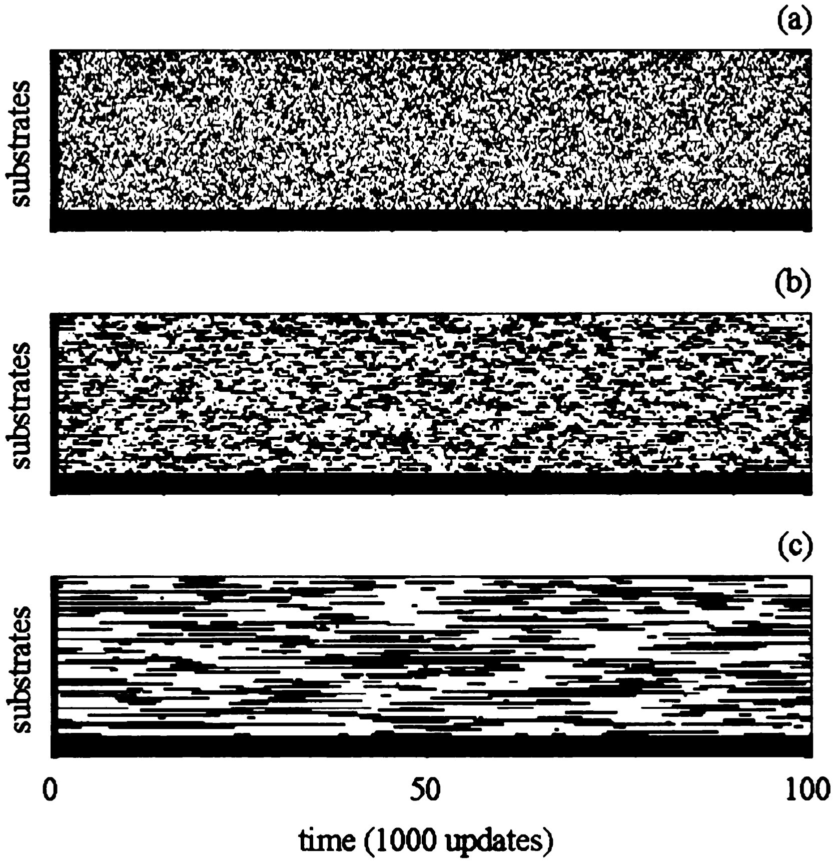


Figure 8. Different periods of change in substrate metabolic values. The 9 fixed ancestral and 68 changing substrates are at various positions on the y-axis. Black indicates that a substrate has a positive metabolic value (nutrient), while white indicates that a substrate has a negative metabolic value (poison) at that particular point during the experiment. Light grey marks the initial period where all changing substrates had a metabolic value of zero. Panels **(a)**, **(b)**, and **(c)** represent 3u, 30u, and 300u periods of environmental change, respectively.

the default instruction set with 25 instructions plus the two divide instructions. Three different types of mutations occur in the system: point, insert, and delete, at rates of 0.002, 0.0005, and 0.0005 per instruction copied, respectively. The carrying capacity is 3,600 organisms and when a population is full, each new birth results in a death of a randomly chosen organism, thus maintaining constant population size. After restarting the populations from Chapter 2 here, we allowed them to acclimate and evolve for 1,000 updates in the constant, familiar environment from Chapter 2 before introducing the additional changing substrates.

RESULTS AND DISCUSSION

The trajectories of the relative abundance of sexual organisms over time showed great diversity across different populations in our study (Figure 9). Some populations started and remained either primarily sexual or asexual, while others switched their mode of reproduction multiple times. Most populations were, at any given moment, homogeneous in terms of their reproductive mode, being comprised of either only sexual or only asexual organisms. The most common exceptions were the rapid transitions between the sexual and asexual types when both coexisted for up to a few dozen generations. Only in a handful of populations did a mixture of sexual and asexual organisms persist for an extended period of time (e.g. Figure 9c). We analyzed the data in two different ways: first, by recording the dominant mode of reproduction at the end of the study; second, by calculating the proportion of time during which each experimental population was predominantly sexual.

We found that sexual reproduction was frequently present and often favored over asexual (Figure 10a). For example, when the environmental change period was 10u and metabolic values were (-1, 5), more than 65% of populations were predominantly sexual at the end of the study. However, because many populations repeatedly switched the mode of reproduction throughout the study, our data may depend on the choice of the final time point. For example, a population that spent the majority of its history as asexual might better be classified as such, even if it evolved to sexual reproduction shortly before the end of the experiment (e.g. Figure 9b). Thus, we also measured the proportion of time during which a population was dominated by sexually reproducing organisms. Fortunately, both the final (top) and time-average (bottom) measures of reproductive mode yield the same general pattern (Figure 10). Populations are likely to be dominated by sexual organisms when substrates are rapidly changing between nutrients and poisons, and when the benefits of metabolizing nutritious substrates are also high. Both effects of the environmental period of change and of the metabolic values are highly significant (two-way ANOVA, $F_{5,936} = 10.153$, $P < 0.001$ for the period of environmental change and $F_{3,936} = 74.983$, $P < 0.001$ for the metabolic values, Table 4), as is the interaction between these two factors ($F_{15,936} = 5.039$, $P < 0.001$, Table 4). In the two environmental regimes where the value of metabolizing a nutrient was less then or equal to that the cost of consuming a poison, (-1, 1) and (-3, 1), asexual reproduction dominated regardless of the period of environmental change. By contrast, in the two regimes where nutrient benefits outweighed the poison cost, (-1, 3) and (-1, 5), sexual reproduction prevailed as often as asexual when substrate metabolic values were changing rapidly, while asexual reproduction prevailed in more slowly changing

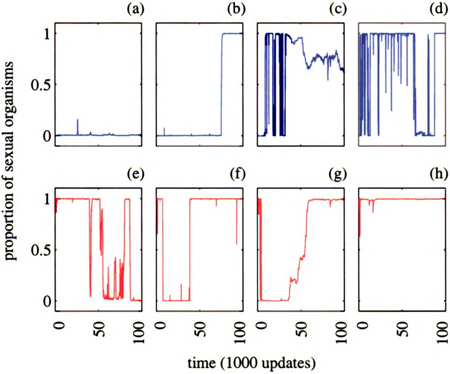
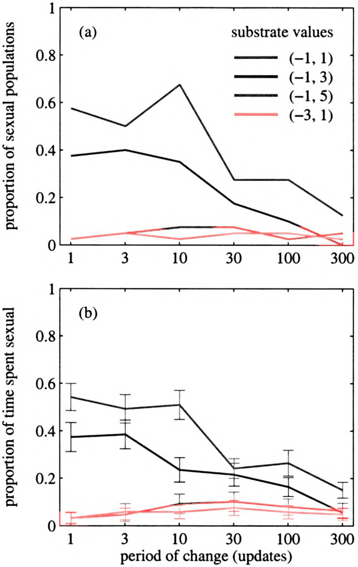


Figure 9. Proportion of sexual organisms present in a population over time. Ancestral populations were asexual for panels **(a)-(d)** and sexual for panels **(e)-(h)**. All the populations evolved at the 3u period of environmental change with (-1, 5) metabolic values. Populations were chosen to illustrate the different types of population trajectories observed in the full data set.

Figure 10. Prevalence of sexual reproduction under 24 different treatments. **(a)** Number of sexual populations that evolved in environments with different periods of environmental change and metabolic values. Populations were categorized as sexual if more than half of the organisms at the end of the experiment reproduced sexually. **(b)** Proportion of time populations spent as sexual during evolution was measured for each population and then averaged over the 40 populations that evolved at each period of environmental change and substrate value combination. Error bars represent one standard error of the mean.


(a)
substrate values
(−1, 1)
(−1, 3)
(−1, 5)
(−3, 1)
proportion of sexual populations
0
0.2
0.4
0.6
0.8
1
1
3
10
30
100
300
(b)
proportion of time spent sexual
0
0.2
0.4
0.6
0.8
1
1
3
10
30
100
300
period of change (updates)

environments. Our results show that sex is beneficial only in environments with certain combinations of the rates and magnitudes of environmental change and the metabolic values, and, moreover, that the interaction between these two factors is also important for the evolution of sex.

To this point, our experiments show that asexual reproduction dominates in slowly changing environments while sexual reproduction prevails – or at least holds its own – when the environmental change is rapid and the benefits of consuming nutrients sufficiently great. We now turn to several subsidiary experiments to examine four related issues. First, we examine whether the initial reproductive mode exerts an influence on the prevalence of sex. Second, we examine whether ancestral features of genetic architecture matter for the evolution of sexual reproduction. Third, we explore whether, at some point, environmental change is so rapid that populations simply cannot adapt fast enough and sex again loses its advantage. Fourth, we ask whether sexually reproducing

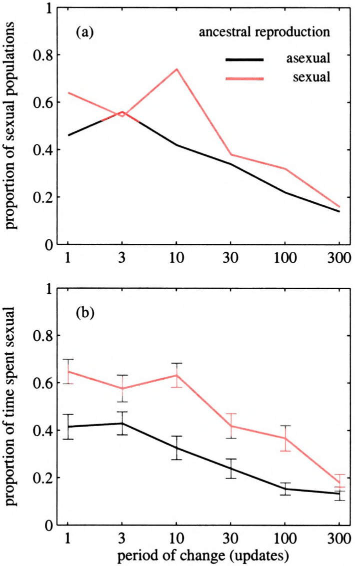
Source of variation	Sum of squares	*df*	Mean square	F	P
Period of change	3.584	5	0.717	10.153	< 0.001
Metabolic values	15.881	3	5.294	74.983	< 0.001
Interaction	5.337	15	0.356	5.039	< 0.001
Error	66.081	936	0.071		

Table 4. Two-way ANOVA testing the effects of the period of environmental change and the substrate metabolic values on the proportion of time populations were dominated by sexual organisms. Using all treatments and data from Figure 10b.

organisms evolve to have higher fitness than asexual ones. In these additional experiments, we focus on the populations that all evolved with the same metabolic values, namely (-1, 5). To increase our statistical power, we generated 60 additional populations (30 initially sexual and 30 initially asexual) at each of the six different periods of environmental change used previously, for a total of 600 populations that experienced the (-1, 5) metabolic values.

We first examined the potential effect of the mode of reproduction of the ancestral population on the prevalence of sex during evolution in changing environments. We found that sexual organisms dominated the populations that started as sexual for a greater proportion of time than the populations that started as asexual (Figure 11, Figure 12b). Sexual populations were also more common overall at the end of the evolution (Figure 12a). The effect of initial mode of reproduction on the time averaged dominant mode of reproduction during evolution in changing environments was highly significant (two-way ANOVA, $F_{1,588} = 5.267$, $P < 0.001$, Table 5), while the interaction between the initial mode of reproduction and the period of environmental change was not significant ($F_{5,588} = 1.783$, $P = 0.114$, Table 5). The lack of a significant interaction reinforces the fact that sex was more common, regardless of ancestral mode, in more rapidly changing environments ($F_{5,588} = 21.207$, $P < 0.001$, Table 5). However, a significant lag in response to selection for a particular reproductive mode could also yield this data. For example, if populations that start as sexual do not immediately switch to asexual reproduction when asexual reproduction is favored, then our data could show a greater proportion of time spent sexual and incorrectly suggest that sex is favored. To address this issue, we reanalyzed the data by calculating the proportion of time that populations

Figure 11. Relationship between the prevalence of sex and the ancestral mode of reproduction. Fifty populations evolved from sexual and 50 from asexual ancestors in environments with (-1, 5) metabolic values. Number of sexual populations **(a)** and proportion of time spent sexual **(b)** were measured as before (see text, Figure 10), and are shown here separately for populations with different ancestral reproductive modes. Error bars represent one standard error of the mean.


(a)
ancestral reproduction
asexual
sexual
proportion of sexual populations
0
0.2
0.4
0.6
0.8
1
1
3
10
30
100
300
(b)
proportion of time spent sexual
0
0.2
0.4
0.6
0.8
1
1
3
10
30
100
300
period of change (updates)

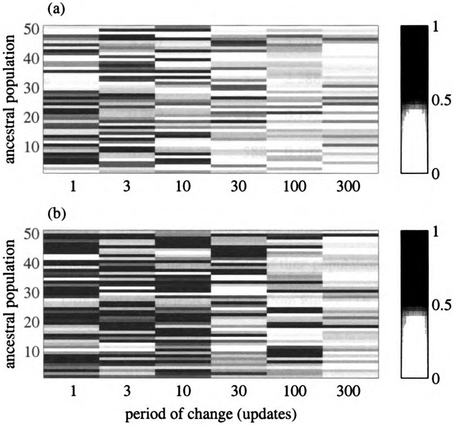


Figure 12. Proportion of time during experimental evolution that each population was predominantly sexual. Dark shades indicate populations that were mostly sexual over time, while light ones indicate populations that were mostly asexual. All populations evolved in environments with metabolic values (-1, 5). Populations with asexual ancestors are presented in panel **(a)** and populations with sexual ancestors in panel **(b)**.

Source of variation	Sum of squares	*df*	Mean square	F	P
Initial reproductive mode	5.267	1	5.267	48.824	< 0.001
Period of change	11.439	5	2.288	21.207	< 0.001
Interaction	0.962	5	0.192	1.783	0.114
Error	63.431	588	0.108		

Table 5. Two-way ANOVA testing the effects of the initial mode of reproduction and period of environmental change on the proportion of time populations were dominated by sexual organisms. Using only treatments and data from Figure 11b.

Source of variation	Sum of squares	*df*	Mean square	F	P
Initial reproductive mode	2.623	1	2.623	17.021	< 0.001
Period of change	14.694	5	2.939	19.071	< 0.001
Interaction	1.143	5	0.229	1.484	0.193
Error	90.608	588	0.154		

Table 6. Two-way ANOVA testing the effects of the initial mode of reproduction and period of environmental change on the proportion of time populations were dominated by sexual organisms. Only the data from the second half of the evolutionary experiments shown in Figure 11b are used (see text).

were sexual only during the second half of the experiments. We found the same qualitative pattern to hold, with significant effects of the initial mode of reproduction and period of change, but no significant interaction between them (Table 6). The overall effect of the initial mode of reproduction was reduced, however, and during the second part of the study, the proportion of time that sexual reproduction was dominant differed by less than 10% between populations started with the two modes of reproduction.

A change in the mode of reproduction is inherently more difficult than simply the maintenance of the existing mode. Moreover, asexual organisms may have evolved genomes that are physically organized in such a way that it is more difficult to evolve sexual reproduction; for example, favorable genetic combinations assembled within asexual genomes may easily be disrupted by recombination. Because our work builds on the evolutionary experiments from Chapter 2, we have an abundance of information about the ancestral populations from which all the populations in this study evolved. Using the data from Chapter 2, we investigate whether the properties of the ancestral populations (besides their mode of reproduction) influence the evolution of sex in changing environments.

The populations in Chapter 2 evolved either strictly sexually or strictly asexually for over 10,000 generations. At the end of that evolution we measured their fitness, genome length, modularity and fitness decay under random mutations. We found that sexual genomes tended to be bigger, on average, and had evolved to be more modular in two different ways. The regions of the genome responsible for metabolizing different nutrients overlapped less (high functional modularity, FM) and were more condensed (high physical modularity, PM) in sexual organisms than in asexual organisms. To determine the effects of random mutations as well as their interactions, we constructed millions of mutants and measured their fitness. We then fit the power function to the data, $\log_{10}W = -\alpha M^{\beta}$, where W is fitness and M the number of random mutations. Parameters α and β reflect robustness to individual mutations and mutational interactions, respectively. We found that sexual organisms were, on average, more robust to mutations ($\alpha_{sex} < \alpha_{asex}$) than asexual organisms. Interactions among mutations were on average antagonistic, but less so in sexual then in asexual organisms ($\beta_{asex} < \beta_{sex} < 1$). There are at least two distinct ways in which genetic architecture in general, and modularity and epistasis in particular, might promote sexual reproduction. First, more modular genomes may allow faster exchange of the modules (metabolic building blocks), making sexual organisms more evolvable (Schlosser and Wagner 2004; Wagner and Altenberg 1996). Second, synergistic epistatic interactions between deleterious mutations may facilitate removal of those mutations through recombination and thus promote sexual reproduction (De Visser and Hoekstra 1998; Kondrashov 1982; Michod and Levin 1988; Wolf et al. 2002). To test whether modularity and epistasis influenced the evolution of sex in the experiments in changing environments, we performed discriminant analysis on the ancestral parameters using the 600 populations evolved with (-1, 5) metabolic values. In particular, for the six treatments with different periods of environmental change, we constructed discriminant functions that would categorize the final populations from this study as sexual or asexual based on their ancestral fitness, genome length, PM, FM, α, and β. Here, the population is categorized as sexual or asexual depending on the most prevalent reproductive mode at the end of the study. The

Period of change	Wilks' lambda	$F_{6,93}$	P	% correct
1u	0.8791	2.1321	0.0568	60
3u	0.9190	1.3666	0.2363	53
10u	0.9111	1.5115	0.1829	56
30u	0.9325	1.1224	0.3554	52
100u	0.9512	0.7947	0.5764	48
300u	0.9352	1.0736	0.3840	53

Table 7. Discriminant analysis for the final mode of reproduction under six periods of environmental change. Each discriminant function was constructed using the ancestral fitness, genome length, PM, FM, α, and β values (see text). The P values indicate whether the function was significant as a whole. The percentages reflect the number of populations that were correctly classified as sexual or asexual using the discriminant function.

six discriminant functions were, on average, able to correctly classify 53.67% of the final populations, but none of the functions were statistically significant (Table 7). When we included the initial mode of reproduction as an additional variable for the discriminant function, on average, 54.83% of populations were correctly classified. Those discriminant functions were significant in only two out of six cases (Table 8). However, both significant discriminant functions assigned the largest weight to the initial mode of reproduction, whose influence was already apparent as described earlier (Figure 11,

Period of change	Wilks' lambda	$F_{6,93}$	P	% correct
1u	0.8600	2.1394	0.0470	63
3u	0.9141	1.2353	0.2917	51
10u	0.8049	3.1848	0.0046	63
30u	0.9316	0.9649	0.4616	50
100u	0.9358	0.9010	0.5094	50
300u	0.9352	0.9108	0.5018	52

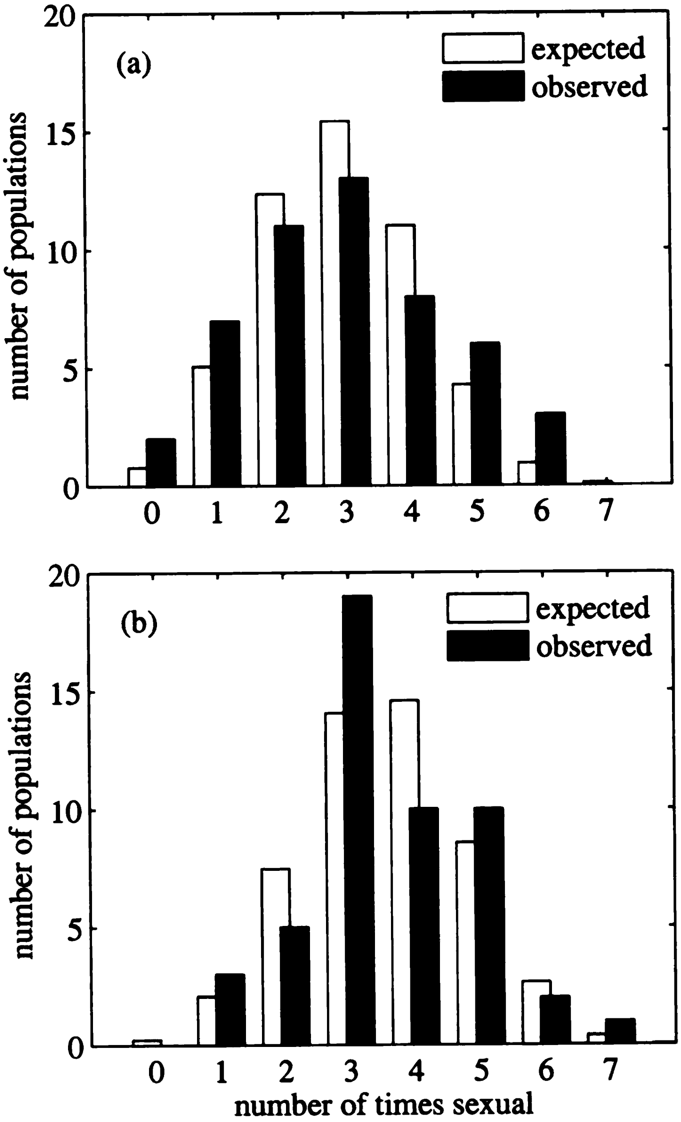
Table 8. Discriminant analysis for the final mode of reproduction under six periods of environmental change. Each discriminant function was constructed using the ancestral mode of reproduction, fitness, genome length, PM, FM, α, and β values (see text). The P values indicate whether the function was significant as a whole. The percentages reflect the number of populations that were correctly classified as sexual or asexual using the discriminant function.

Figure 12). We conclude that genetic architecture had little effect on the subsequent propensity to evolve or maintain sexual reproduction in changing environments.

Further support for this essentially negative result came from the analysis of the number of times sex evolved from each starting population. The same ancestral at each period of environmental change. If a certain population was strongly predisposed to evolve or maintain sexual reproduction, it might be sexual in a disproportionate number of those 6 experiments. Here, we used the predominant final mode of reproduction to categorize a population as sexual or asexual. We first counted the number of times each starting population evolved to be sexual and categorized the populations accordingly (observed values, Figure 13). There were 7 categories, corresponding to ancestral populations that evolved (or remained) sexual 0 through 6 times. We then preformed a randomization test by comparing the observed distribution with a random null-distribution. In generating the null-distribution, we started by randomly selecting a hypothetical collection of sexual populations for each period of environmental change. The number of populations in the collections was determined by the actual number of populations that were dominated by sexual organisms at the end of our experiments. We then counted the number of times each population was in collections from different periods of change. This was repeated 1000 times and the average number of populations in each category was the random distribution (expected values, Figure 13). Finally, we compared the random and the observed distributions, treating populations that evolved from sexual and asexual separately. There was no statistically significant difference between the observed and randomly generated distributions in either case (Chi-square = 10.855, df = 6, P = 0.0930 for asexual ancestors; Chi-square = 4.339, df = 6, P = 0.6309 for sexual ancestors). This result supports the interpretation that the general properties of the ancestral populations do not strongly affect the outcome of evolution of sexual reproduction in changing environments.

Contrary to our expectations, sexual reproduction dominated in the experiments with energy values of (-1, 5) even at the shortest period of environmental change, with one substrate changing from a nutrient to a poison and another from a poison to a nutrient at every update (Figure 11). A closer consideration of our environmental setup suggests that, while the environment may extremely variable, nutrients must also reoccur very

Figure 13. Number of times each ancestral population evolved to be sexual. Populations evolved in environments with metabolic values (-1, 5). Populations with asexual ancestors are presented in panel **(a)**, and populations with sexual ancestors in panel **(b)**. Black bars represent the observed number of times individual ancestral populations became or remained sexual after evolution in environments with 6 different periods of change. White bars represent the expected values based on the randomization procedure described in the text.


(a)
expected
observed
number of populations
0
5
10
15
20
0
1
2
3
4
5
6
7
(b)
expected
observed
number of populations
0
5
10
15
20
0
1
2
3
4
5
6
7
number of times sexual

frequently. For example, when the period of environmental change is 1 update, each substrate on average oscillates though periods of 25 updates (~3 generations) as a nutrient, followed by 68 updates (~8 generations) as a poison. As we have a finite number of substrates available in the system, we cannot create environments that will continuously present the evolving populations with completely novel nutrients and poisons. However, we can lower the number of nutrients present at one time, and thus extend the time between occurrences of individual nutrients in the environment. We conducted additional experiments at the metabolic values (-1, 5) with only 10 of the 68 variable substrates as nutrients at any given time, in contrast to the 25 present as nutrients in all our previous experiments, with all else equal. The environment was changing in the same way as in the previous experiments and 20 initially sexual and 20 initially asexual populations evolved at every 1u, 3u, 10u, 30u, 100u, and 300u period of change. We find now that the asexual reproduction dominated at all periods of environmental change (Figure 14). We contrast these experiments to those in which the 40 populations evolved with identical metabolic values (-1, 5), but with 25 nutrients present at a time (Figure 10). We find that in the 25-nutrient experiments the proportion of time spent as sexual continues upward (or perhaps levels off) at the shortest periods of environmental change, while in the 10-nutrient experiments it clearly peaks at the intermediate values and then declines as the environment changes even faster. The period of environmental change had a significant effect on the reproductive mode in the 10-nutrient experiments (one-way ANOVA, $F_{5,234} = 2.964$, $P = 0.013$, Table 9). We fit linear and a quadratic function to both sets of data to characterize the shape of the observed relationship between the period of change and prevalence of sex. Given that the period of change

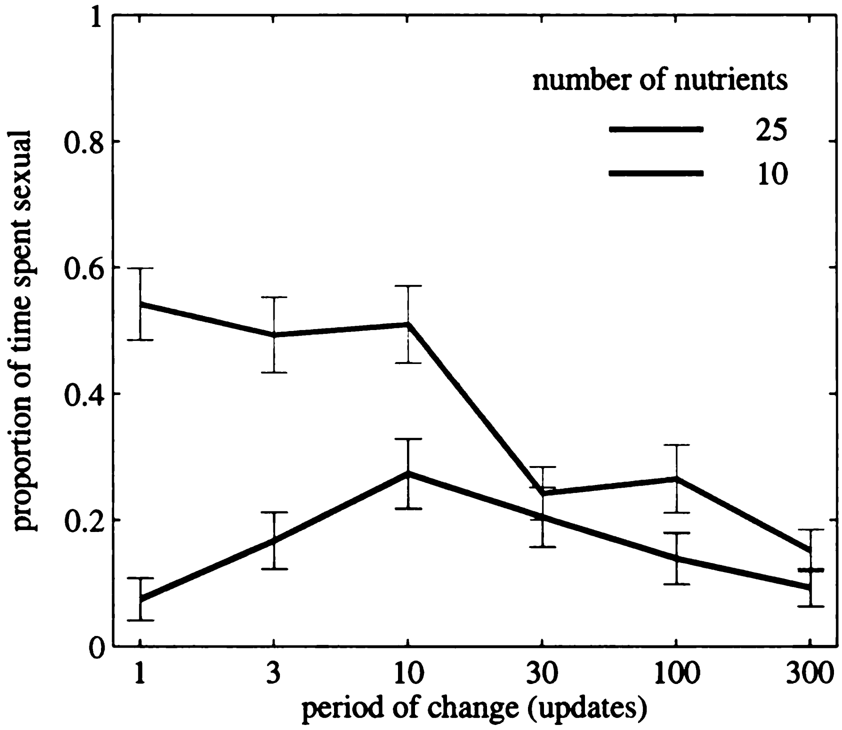


Figure 14. Prevalence of sexual reproduction in environments with different numbers of nutrients. All populations evolved in the environments with (-1, 5) metabolic values. Populations had either 25 or 10 nutrients present any the time among the 68 changing substrates. Data for populations with 25 nutrients are reproduced from Figure 10 for comparison. Proportion of time spent sexual was measured as before (see text, Figure 10). Error bars represent one standard error of the mean.

spans three orders of magnitude, we $\log_{10}$ transformed the values for the period of change prior to the regression analysis. In 10-nutrient environments, linear function was not significant, while the quadratic function was highly significant (Table 10). In contrast, in the 25-nutrient populations, the linear function was significant, but the quadratic term did not significantly reduce the unexplained variation (Table 10). Our results suggest that while sex may thrive in rapidly changing environments, there can also be too much change under certain scenarios, in which case asexual reproduction prevails again.

Finally, we examine the fitness values of sexual and asexual organisms evolved in the 600 populations with metabolic values (-1,5). We identify the periods of time during the evolution of each population when the majority of organisms were sexual, and we calculate the mean of the fitness values recorded during those times. (Because the environments are constantly changing while maintaining the same number of resources and poisons, averaging fitness values across time is reasonable.) The same calculation is done for the periods when a population was primarily reproducing asexually (Figure 15). The mode of reproduction did not have a significant major effect on fitness (two-way ANOVA, $F_{1,1070} = 2.102$, $P = 0.147$, Table 11), but the period of environmental change did ($F_{5,1070} = 2.555$, $P = 0.026$, Table 11). The interaction between the mode of reproduction and the period of the environmental change was also marginally significant ($F_{5,1070} = 1.974$, $P = 0.080$, Table 11). The general trend in relative fitness of sexual and asexual populations confirms our previous results. In rapidly changing environments, sexual organisms were favored over asexual ones, and the mean fitness of sexual

Source of variation	Sum of squares	*df*	Mean square	F	P
Period of change	1.081	5	0.216	2.964	0.013
Error	17.070	234	0.073		

Table 9. One-way ANOVA testing the effects of the period of environmental change on the proportion of time populations were dominated by sexual organisms. Using data from the 10-nutrient treatment shown in Figure 14.

Number of nutrients	Type of function	R^2	$F_{1,596}$	P
10	Linear	< 0.001		
	Quadratic	0.051	31.368	< 0.001
25	Linear	0.154		
	Quadratic	0.156	1.412	0.235

Table 10. Linear and quadratic fits of proportion of time spent sexual to $\log_{10}$ transformed period of environmental change for populations evolving with 10 or 25 nutrients at a time. We conducted the Partial F test for significance of the marginal improvement associated with the addition of the quadratic term to the model. We show the R^2 values for both linear and quadratic fit, the Partial F statistic and the P value for the Partial F test.

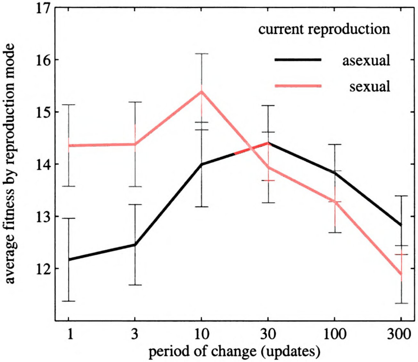


Figure 15. Average fitness of predominantly sexual and predominantly asexual populations. Fitness values that were recorded when a population was predominantly sexual were averaged, and the same was done for the fitness values from the times when a population was predominantly asexual. The fitness while sexual or asexual was then averaged over the 100 populations that evolved at each period of environmental change. All populations evolved in the environments with (-1, 5) metabolic values. Error bars represent one standard error of the mean.

Source of variation	Sum of squares	*df*	Mean square	F	P
Initial reproductive mode	93.575	1	93.575	2.102	0.147
Period of change	568.782	5	113.756	2.555	0.026
Interaction	439.337	5	87.867	1.974	0.080
Error	47630.451	1070	44.514		

Table 11. Two-way ANOVA testing the effects of the mode of reproduction and the period of environmental change on mean fitness. Using only treatments and data from Figure 15.

populations was higher than that of asexual populations. Conversely, in more slowly changing environments (30u, 100u, 300u), asexual organisms dominated the populations most of the time and the mean fitness of asexual populations was slightly higher than that of the sexual populations.

Over the past century, much thought has gone toward understanding the evolution sexual reproduction (Bell 1982; Michod and Levin 1988; Weismann 1889), but experiments in the area difficult and rare (Rice 2002). Here, we use digital organisms to experimentally investigate the conditions in which sex can persist or even invade when sexual and asexual organisms directly compete with each other. The potential for sex to evolve in changing environments has been theoretically established in many studies, but not experimentally tested in the past (Otto and Michalakis 1998). We find that sex can prevail when populations evolve in rapidly changing environments. The maintenance and invasion of sexual reproduction were only weakly affected by the ancestral mode of reproduction, and other ancestral properties, such as genome length and genetic architecture, also had little effect on evolution of sexual reproduction in changing environments. Our experiments cannot say whether changing environments are the only or the most important factor promoting the evolution of sex in the natural world (West et al. 1999), but our study does demonstrate experimentally that environmental factors and their interactions can play a very significant role in the evolution and maintenance of sex.

APPENDIX

PHENOTYPES IN AVIDA

As described in the main text, when an organism performs one of nine basic logic operations on one or two random 32-bit strings and then outputs the bitwise-correct result, it obtains additional energy that accelerates the execution of its genomic program. The logic rules for all the operations are presented in Table 12 below. For example, if bit A = 1 and bit B = 0, then (A OR B) = 1. These rules are defined on single-bit inputs; for an organism to be rewarded for any operation, it must correctly perform it on all 32 bits of the number strings.

Consider an organism that obtained the following two inputs and then output the string below.

Input A: 01010101110000001110101010101100

Input B: 10000110101000111101010110011110

Output : 11010111111000111111111110111110

The organism would receive the energy reward for performing the OR operation, because it correctly calculated the OR function for all 32 pairs of the corresponding bits in Input A and Input B and output the correct result.

The ability to perform a logic operation is scored as a phenotypic trait in the GP maps. Figure 3c (main text) shows two different organisms that can perform eight of the nine rewarded operations. In these examples, both the sexual and asexual organisms have evolved the ability to perform the NOT, NAND, AND, ORN, OR, ANDN, NOR, and EQU operations, but neither one can perform the XOR operation that is represented by the penultimate column.

Input		Logic operation								
A	B	NOT	NAND	AND	ORN	OR	ANDN	NOR	XOR	EQU
0	0	1	1	0	1	0	0	1	0	1
0	1	1	1	0	0	1	0	0	1	0
1	0	0	1	0	1	1	1	0	1	0
1	1	0	0	1	1	1	0	0	0	1

Table 12. Truth table for nine logic operations. NOT is preformed only on Input A, while the other eight operations are functions of both inputs

RESOLVING GENOME-LENGTH ISSUES

In the main text, we reported that sexual populations evolved larger genomes, on average, than did asexual populations. This difference was driven by a subset of populations that had evolved very large genomes. Out of the 100 populations with each reproductive mode, 35 sexual populations evolved genomes that averaged at least twice the ancestral length of 50 instructions, while only 9 asexual populations did so. These very large genomes typically evolved via genome doublings, and sexual reproduction in Avida evidently increased the rate of genome doubling events, created more favorable genetic combinations in larger genomes, or both.

To determine whether genome doublings and the resulting difference in average genome length between evolved sexual and asexual populations might bias or confound the comparisons of their genetic architectures, we performed two supplementary analyses described below. The first analysis excluded all those populations that evolved average genome lengths twice or more the ancestral length. The second analysis used 100 additional populations with each reproductive mode, in which genome doublings were prevented during the experiment itself.

Both supplementary analyses eliminated the greater average genome length of the sexual populations and, in fact, they reversed the difference such that the asexual populations had on average significantly longer genomes. Both also eliminated the higher mean fitness values of the sexual populations relative to the asexual populations. However, both supplementary analyses confirmed all the differences in genetic architecture between sexual and asexual populations: (i) sexual genomes were more modular (higher *PM* and *FM*); (ii) sexual populations were more robust to the effects of single mutations (lower α); and (iii) asexual populations had stronger net directional epistasis tending to alleviating interactions (lower β).

REDUCED DATA SET

We excluded all populations with average genome length $\geq$100 instructions, which is twice the ancestral length. We repeated the original analyses on this reduced data set, which included 65 sexual and 91 asexual populations. The differences between sexual and asexual populations in *PM, FM*, α, and β remained significant and in the same direction as in the full data set (Table 13). However, average genome length was now significantly greater in the asexual organisms, the opposite outcome to the full data set, so that genome length cannot be driving these differences in genetic architecture.

GENOME DOUBLINGS PREVENTED

Analysis of the reduced data set shows that the larger tail of long genomes in the sexual populations was not responsible for the evolved differences in genetic architecture between sexual and asexual organisms. We sought further confirmation by evolving an additional 200 populations with the experimental conditions identical to the original runs in all but one respect: genome doublings were prevented from ever occurring in these experiments by imposing a 10% limit on the difference in genome length between parents and offspring. Whenever a divide instruction (sexual or asexual) was about to be executed, the Avida program checked the genome length of the incipient offspring. If the potential genome was more than 10% longer (or shorter) than its parent's genome, the divide instruction was skipped and the execution proceeded to the next instruction in the genome. Thus, no offspring were ever produced with genome doublings or other radical changes in genome length. Very large genomes did not evolve in these modified runs and, in fact, sexual populations had significantly shorter genomes, on average, than did asexual populations. However, the genetic architectural features *PM*, α, and β differed significantly and in the same direction as observed in the original experiments (Table 14). The difference in *FM* was no longer significant but remained in the same direction.

response variable	mean sexual (± s.d.)	mean asexual (± s.d.)	P
$\log_{10}$ fitness	4.835 (1.230)	5.033 (1.104)	0.303
$\log_{10}$ genome length	1.752 (0.100)	1.826 (0.059)	< 0.001
physical modularity, *PM*	0.775 (0.057)	0.689 (0.066)	< 0.001
functional modularity, *FM*	0.803 (0.084)	0.761 (0.077)	0.002
average effect of single point mutations on fitness, α	0.724 (0.173)	0.792 (0.176)	0.019
net directional epistasis, β	0.904 (0.051)	0.854 (0.088)	< 0.001

Table 13. Comparisons of properties between sexual and asexual evolved populations in the reduced data set. The P values are based on two-tailed *t*-tests.

response variable	mean sexual (± s.d.)	mean asexual (± s.d.)	P
$\log_{10}$ fitness	4.678 (1.107)	4.944 (1.242)	0.112
$\log_{10}$ genome length	1.703 (0.013)	1.845 (0.061)	< 0.001
physical modularity, *PM*	0.717 (0.064)	0.694 (0.073)	0.021
functional modularity, *FM*	0.771 (0.078)	0.761 (0.080)	0.356
average effect of single point mutations on fitness, α	0.765 (0.162)	0.827 (0.183)	0.004
net directional epistasis, β	0.901 (0.056)	0.828 (0.107)	< 0.001

Table 14. Comparisons of properties between sexual and asexual evolved populations with genome doublings prevented. The P values are based on two-tailed *t*-tests.

REFERENCES

Adami, C. 1998. *Introduction to Artificial Life*. Springer, New York.

Adami, C. 2002. Ab initio modeling of ecosystems with artificial life. *Natural Resource Modeling* **15**:133-145.

Adami, C., C. Ofria, and T. C. Collier. 2000. Evolution of biological complexity. *Proceedings of the National Academy of Sciences of the United States of America* **97**:4463-4468.

Andersson, D. I., and D. Hughes. 1996. Muller's ratchet decreases fitness of a DNA-based microbe. *Proceedings of the National Academy of Sciences of the United States of America* **93**:906-907.

Barton, N. H. 1995. A general model for the evolution of recombination. *Genetical Research* **65**:123-144.

Beldade, P., K. Koops, and P. M. Brakefield. 2002. Modularity, individuality, and evo-devo in butterfly wings. *Proceedings of the National Academy of Sciences of the United States of America* **99**:14262-14267.

Bell, G. 1982. *The Masterpiece of Nature*. University of California Press, Berkeley.

Bell, G. 1989. *Sex and Death in Protozoa*. Cambridge University Press, Cambridge.

Chao, L. 1988. Evolution of sex in RNA viruses. *Journal of Theoretical Biology* **133**:99-112.

Chao, L. 1990. Fitness of RNA virus decreased by Muller's Ratchet. *Nature* **348**:454-455.

Chao, L., and T. T. Tran. 1997. The advantage of sex in the RNA virus $\phi 6$. *Genetics* **147**:953-959.

Charlesworth, B. 1993. Directional selection and the evolution of sex and recombination. *Genetical Research* **61**:205-224.

Chow, S. S., C. O. Wilke, C. Ofria, R. E. Lenski, and C. Adami. 2004. Adaptive radiation from resource competition in digital organisms. *Science* **305**:84-86.

De Visser, J., and R. F. Hoekstra. 1998. Synergistic epistasis between loci affecting fitness: evidence in plants and fungi. *Genetical Research* **71**:39-49.

De Visser, J., R. F. Hoekstra, and H. van den Ende. 1997. Test of interaction between genetic markers that affect fitness in *Aspergillus niger*. *Evolution* **51**:1499-1505.

Duarte, E., D. Clarke, A. Moya, E. Domingo, and J. Holland. 1992. Rapid fitness losses in mammalian RNA virus clones due to Muller's ratchet. *Proceedings of the National Academy of Sciences of the United States of America* **89**:6015-6019.

Earl, D. J., and M. W. Deem. 2004. Evolvability is a selectable trait. *Proceedings of the National Academy of Sciences of the United States of America* **101**:11531-11536.

Elena, S. F., and R. E. Lenski. 1997. Test of synergistic interactions among deleterious mutations in bacteria. *Nature* **390**:395-398.

Felsenstein, J. 1974. The evolutionary advantage of recombination. *Genetics* **78**:737-756.

Ferrier, D. E. K., and P. W. H. Holland. 2001. Ancient origin of the Hox gene cluster. *Nature Reviews Genetics* **2**:33-38.

Ghiselin, M. T. 1988. The evolution of sex: A history of competing points of view, pp 7-23 *in* R. E. Michod, and B. Levin, eds. *The Evolution of Sex*. Sinauer, Sunderland, Massachusetts.

Goddard, M. R., H. Charles, J. Godfray, and A. Burt. 2005. Sex increases the efficacy of natural selection in experimental yeast populations. *Nature* **434**:636-640.

Hartwell, L. H., J. J. Hopfield, S. Leibler, and A. W. Murray. 1999. From molecular to modular cell biology. *Nature* **402**:C47-C52.

Kirschner, M., and J. Gerhart. 1998. Evolvability. *Proceedings of the National Academy of Sciences of the United States of America* **95**:8420-8427.

Kondrashov, A. S. 1982. Selection against harmful mutations in large sexual and asexual populations. *Genetical Research* **40**:325-332.

Kondrashov, A. S. 1993. Classification of hypotheses on the advantage of amphimixis. *Journal of Heredity* **84**:372-387.

Kondrashov, A. S., and L. Y. Yampolsky. 1996. Evolution of amphimixis and recombination under fluctuating selection in one and many traits. *Genetical Research* **68**:165-173.

Laumanns, M., E. Zitzler, and T. Lothar. 2001, On the effects of archiving, elitism, and density based selection in evolutionary multi-objective optimization, pp 181-195 *in* E. Zitzler, K. Deb, L. Thiele, C. A. C. Coello, and D. Corne, eds. *Evolutionary Multi-Criterion Optimization: First International Conference Proceedings*. Springer-Verlag, Heidelberg.

Lawrence, J. G., and J. R. Roth. 1996. Selfish operons: horizontal transfer may drive the evolution of gene clusters. *Genetics* **143**:1843-1860.

Lenski, R. E. 1999. A distinction between the origin and maintenance of sex. *Journal of Evolutionary Biology* **12**:1034-1035.

Lenski, R. E., C. Ofria, T. C. Collier, and C. Adami. 1999. Genome complexity, robustness and genetic interactions in digital organisms. *Nature* **400**:661-664.

Lenski, R. E., C. Ofria, R. T. Pennock, and C. Adami. 2003. The evolutionary origin of complex features. *Nature* **423**:139-144.

Lewis, W. M. 1983. Interruption of synthesis as a cost of sex in small organisms. *American Naturalist* **121**:825-834.

Lively, C. M. 1987. Evidence from a New Zealand snail for the maintenance of sex by parasitism. *Nature* **328**:519-521.

Lively, C. M., and M. F. Dybdahl. 2000. Parasites adaptation to locally common host genotypes. *Nature* **405**:679-681.

Lynch, M., R. Bürger, D. Butcher, and W. Gabriel. 1993. The mutational meltdown in asexual populations. *Journal of Heredity* **84**:339-344.

Lynch, M., J. Conery, and R. Bürger. 1995. Mutation accumulation and the extinction of small populations. *American Naturalist* **146**:489-518.

Malmberg, R. L. 1977. Evolution of epistasis and advantage of recombination in populations of bacteriophage T4. *Genetics* **86**:607-621.

Maynard Smith, J. 1971. What use is sex? *Journal of Theoretical Biology* **30**:319-335.

Maynard Smith, J. 1978. *The Evolution of Sex*. Cambridge University Press, Cambridge.

Maynard Smith, J. 1988. The evolution of recombination, pp 106-125 *in* R. E. Michod, and B. Levin, eds. *The Evolution of Sex*. Sinauer, Sunderland, Massachusetts.

McPhee, C. P., and A. Robertson. 1970. The effect of supressing crossing-over on the response to selection in *Drosophila melanogaster*. *Genetical Research* **16**:1-16.

Michod, R. E., and B. Levin, eds. 1988. *The Evolution of Sex*. Sinauer, Sunderland, Massachusetts.

Misevic, D., C. Ofria, and R. E. Lenski. 2004. Sexual reproduction and Muller's ratchet in digital organisms, pp 340-345 *in* J. B. Pollack, M. Bedau, P. Husbands, T.

Ikegami, and R. A. Watson, eds. *Proceedings of Artificial Life IX.* MIT Press, Cambridge, Massachusetts

Misevic, D., C. Ofria, and R. E. Lenski. 2006. Sexual reproduction reshapes the genetic architecture of digital organisms. *Proceedings of the Royal Society of London Series B-Biological Sciences* **273**:457-464.

Muller, H. G. 1964. The relation of recombination to mutational advance. *Mutation Research* **1**:2-9.

Nowak, M., and P. Schuster. 1989. Error thresholds of replication in finite populations mutation frequencies and the onset of Muller's ratchet. *Journal of Theoretical Biology* **137**:375-395.

Ofria, C., and C. O. Wilke. 2004. Avida: a software platform for research in computational evolutionary biology. *Artificial Life* **10**:191-229.

Otto, S. P., and Y. Michalakis. 1998. The evolution of recombination in changing environments. *Trends in Ecology & Evolution* **13**:145-151.

Pál, C., and L. D. Hurst. 2003. Evidence for co-evolution of gene order and recombination rate. *Nature Genetics* **33**:392-395.

Pál, C., and L. D. Hurst. 2004. Evidence against the selfish operon theory. *Trends in Genetics* **20**:232-234.

Pürgel-Bennett, A. 1997. Modelling evolving populations. *Journal of Theoretical Biology* **185**:81-95.

Ravasz, E., A. L. Somera, D. A. Mongru, Z. N. Oltvai, and A. L. Barbasi. 2002. Hierarchical organization of modularity in metabolic networks. *Science* **297**:1551-1555.

Rice, W. R. 2002. Experimental tests of the adaptive significance of sexual recombination. *Nature Reviews Genetics* **3**:241-251.

Rice, W. R., and A. K. Chippindale. 2001. Sexual recombination and the power of natural selection. *Science* **294**:555-559.

Sasaki, A., and Y. Iwasa. 1987. Optimal recombination rate in fluctuating environments. *Genetics* **115**:377-388.

Schlosser, G., and G. P. Wagner, eds. 2004. *Modularity in Development and Evolution.* University of Chicago Press, Chicago.

Sokal, R. R., and F. J. Rohlf. 1995. *Biometry.* Freeman, New York.

Wagner, G. P. 1996. Homologues, natural kinds and the evolution of modularity. *American Zoologist* **36**:36-43.

Wagner, G. P., and L. Altenberg. 1996. Complex adaptations and the evolution of evolvability. *Evolution* **50**:967-976.

Wagner, G. P., M. D. Laubichler, and H. Bagheri-Chaichian. 1998. Genetic measurement theory of epistatic effects. *Genetica* **103**:569-580.

Waxman, D., and J. R. Peck. 1999. Sex and adaptation in a changing environment. *Genetics* **153**:1041-1053.

Weismann, A. 1889. *Essays upon Heredity and Kindred Biological Problems.* Clarendon Press, Oxford.

West, S. A., C. M. Lively, and A. F. Read. 1999. A pluralist approach to sex and recombination. *Journal of Evolutionary Biology* **12**:1003-1012.

Wilke, C. O., and C. Adami. 2001. Interaction between directional epistasis and average mutational effects. *Proceedings of the Royal Society of London Series B-Biological Sciences* **268**:1469-1474.

Wilke, C. O., and C. Adami. 2002. The biology of digital organisms. *Trends in Ecology & Evolution* **17**:528-532.

Wilke, C. O., R. E. Lenski, and C. Adami. 2003. Compensatory mutations cause excess of antagonistic epistasis in RNA secondary structure folding. *BMC Evolutionary Biology* **3** (doi:10.1186/1471-2148-3-3).

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

Williams, G. C. 1975. *Sex and Evolution.* Princeton University Press, Princeton.

Winther, R. G. 2001. Varieties of modules: kinds, levels, origins, and behaviors. *Journal of Experimental Zoology* **291**:116-129.

Wolf, H. G., K. Wohrmann, and J. Tomuik. 1987. Experimental evidence for the adaptive value of sexual reproduction. *Genetica* **72**:151-159.

Wolf, J. B., E. D. Brodie, and M. J. Wade, eds. 2002. *Epistasis and the Evolutionary Process.* Oxford University Press, Oxford.

Zitzler, E., K. Deb, and L. Thiele. 2000. Comparison of multiobjective evolutionary algorithms: empirical results. *Evolutionary Computation* **8**:173-195.