DOI: http://dx.doi.org/10.7551/978-0-262-31050-5-ch005

What does sex have to do with it: tracking the fate of deleterious mutations in sexual populations.

Arthur W. Covert III[1,2]*, Lane Smith[1,2], Dakota Z. Derrberry[1,2], and Claus O. Wilke[1,2]

[1] Section of Integrative Biology, Center for Computational Biology and Bioinformatics, and Institute for Cellular and Molecular Biology, The University of Texas at Austin.
[2] BEACON, Center for the Study of Evolution in Action,
*Author to whom correspondence should be sent
Art.Covert@austin.utexas.edu

Abstract

Deleterious mutations sometimes revert into beneficial mutations via epistatic interactions with subsequent mutations. This type of interaction among mutations is called "sign-epistasis." Recent works have explored the role of sign-epistasis in the evolution of asexual populations. Some have indicated that the fixation of sign-epistasic deleterious mutations may be critical for adaptive evolution. However, sign-epistasis is considered to be important only for asexual populations, because recombination in sexual populations tends to disrupts linkage between epistatically interacting mutations. Here, we tested the hypothesis that recombination prevents adaptation via sign-epistatic fitness reversions, by examining deleterious mutations in sexually-reproducing digital organisms. We examined every deleterious mutation that arose on the genealogy between the original ancestor and the final dominant genotype (the "graph of descent"). We show that sign-epistatic pairs of mutations emerged in several replicate populations, and that they contributed positively to the long-term adaptation of the population.

Introduction

Deleterious mutations are generally thought of as a drag on adaptive ability. In rare cases a deleterious mutation may be joined by a compensatory adaptation that ameliorates the deleterious effect. If the compensatory adaptation itself was deleterious in the absence of the original deleterious mutation, then the pair of mutations are individually deleterious, but jointly beneficial. In finite populations these pairs of mutations may then sweep to fixation jointly, rather then having to fix sequentially (Iwasa et al 2004, Weissman et al 2009).

Interactions between mutations that alter their cumulative fitness effect are called epistatic interactions. The most extreme form of epistasis, a change in the fitness effect from deleterious to beneficial, is called sign-epistasis (Weinreich and Chao, 2005).

Theoretical works examining sign-epistasis have found that deleterious mutations in asexual populations can segregate at low frequencies and occasionally be compensated by subsequent mutations. Computational simulations have found actual examples of sign-epistatsis contributing to adaptation (Lenski et al 2003, Cowperthwaite et al 2006, Covert 2010). More recently, an example of a sign-epistatic interaction has been found in *Saccharomyces cerevisiae*, the first known discovery of a sing-epistatic fitness reversal in an organic system (Kvitek and Sherlock 2011).

All works on sign-epistasis to date have considered asexual systems. Presumably, recombination would disrupt any useful sign-epistasic paring, unless the interacting mutations were very tightly linked (Weinreich and Chao 2005). We test this assumption using digital organisms that undergo recombination. We compare two experimental treatments, one in which organisms can suffer deleterious mutations and one in which deleterious mutations are prevented from occurring. We see that populations that do not experience deleterious mutations evolve to significantly lower fitness than populations that do. To identify where sign-epistatic mutations occurred and what effect they had, we reconstruct a complete genealogy ("graph of descent") from the original ancestor of the population to the final, most abundant genotype. We use the graph of descent to isolate specific examples sign-epistatic interactions and examine when they emerged, when they recovered, and what magnitude their epistatic effects had.

Methods

Experimental system. We used the digital-life platform Avida (version 2.12.2) for all experiments. The Avida world holds a population of digital organisms. Digital organisms are self-replicating and evolving computer programs, written in a special-purpose programming language and executed on a grid of virtual CPUs. The computer program defining a digital organism is considered to be the organism's genome. Mutations are random changes in the genome. Here, we only used point mutations, which replace one instruction in the genome with a randomly chosen instruction.

Digital organisms are rewarded with additional energy (CPU time) for the successful computation of logical functions. Thus, there is a selective pressure for digital organisms to evolve the capability to efficiently compute multiple logical functions. Here, we used the standard Avida "logic-9" environment as described in (Lenski et al. 1999, Lenski et al. 2003). This environment rewards one- and two-input logical functions; reward amounts increase with difficulty of the logical function to compute.

Adaptation experiment. We adapted replicate populations of digital organisms for 250,000 updates[1]. Populations were seeded with an organism whose genome consisted of 50 instructions. The seed organism could self-replicate but not perform any of the logical functions. Populations were seeded with a single digital organism; population size rapidly grew to a maximum carrying capacity of 10,000, at which it remained for the remainder of the adaptations. Organisms had a 25% chance of experiencing a single point mutation on divide (replication). This mutation rate translates into a 0.5% probability of mutation per site per generation.

Recombination was implemented as follows: When an organism divided, its offspring was placed into a birth chamber. Organisms remained in the birth chamber until they were joined by another organism, with which they recombined. Recombination occurred at a single cross-over point in the middle of the genome. Recombination between identical genotypes was allowed (Misevic et al 2006).

We ran 50 replicates each of two separate treatments, *Control* and *Revert Deleterious* (RvD). In the Control treatment, all newly divided organisms were replaced in an isolated test environment before they entered the birth chamber. Each organism was tested to determine if it could self-replicate without altering its genome. Those organisms that could self-replicate stably were placed in the birth chamber. Organisms that could not self-replicate stably were sterilized and removed from the population.

In the RvD treatment, we tested organisms for stable self-replication as well as for the presence of a deleterious mutation. If an organism experienced a deleterious (but not lethal) mutation, we reverted the organism's genotype to the parent's genotype[2]. The RvD treatment prevented the occurrence of deleterious point mutations. However, note that recombination could nevertheless create combinations of mutations that were deleterious, even if the individual mutations were not deleterious in their parent organisms.

We used a structured population, to make constructing the graph of descent more tractable. Each population was divided into 100 subpopulations of equal size. Normally when an organism leaves the birth chamber it is placed next to the last parent to contribute a genome to the birth chamber. When organisms left the birth chamber in our structured environment they had a 1 in 20,000 chance of migrating to a new subpopulation (approximately one migration event every other generation). The structured population limits the total number of genomes that may contribute to the final population. Since Avida saves only those genotypes related to the most abundant genotypes still alive in the population, the choice of a structured population made both the file size and the number of genotypes to examine more manageable.

For each evolved population, we measured the fitness of the dominant (most abundant) genotype at the end of 250,000 updates of adaptation. We used this fitness measure to assess long-term evolutionary success of evolved populations.

Construction of the graph of descent. We reconstructed the genealogies of 20 final dominant genotypes (FDGs) from the first 20 replicates in the experiment. Each genealogy included all parent-offspring relationships between the original ancestor and the FDG. We began by identifying the two parents of the FDG. We next identified the parents of the FDG's two parents. Then we identified the parents of the FDG's parent's parents. We continued until we had traced back the entire ancestry to the original ancestor. This procedure resulted in a graph of genotypes spanning from the original common ancestor to the FDG, with bidirectional edges from parent to offspring genotypes.

We modified the Avida software to output the mutations and crossover points for each genotype. Starting with the offspring of the common ancestor, we examined the fitness effect of every mutation on the graph of descent. For each deleterious mutation on the graph of descent we created a "mutation subgraph" that traced the fate of a single mutation. Each deleterious mutation was tracked from its entrance in the graph until either all genotypes on the graph had mutated away from the deleterious mutation or one of the genotypes had undergone a sign-epistatic fitness recovery.

Identification of sign-epistatic mutations on the graph of descent. We used the following algorithm to identify sign-epistatic mutations: We iterate over all deleterious mutations on the graph of descent. For each mutation, we confirm its deleterious fitness effect by undoing only the mutation, not the recombination. We undo or revert a mutation by replacing the mutated instruction with the instruction at the same locus in the parent of the origin. If the fitness of the reverted offspring is less than the fitness of at least one of the parents, we know the mutation is deleterious. We call the first genotype that contains a deleterious mutation the origin (Figure 1). For each offspring containing the deleterious mutation, we similarly revert the mutation to the instruction at the same locus in the parent of the origin. If the reversion increases fitness we know that the mutation is still deleterious in the current genetic background. If the reversion decreases fitness on the current background (or does not alter it), we know that the previously deleterious mutation has undergone a sign-epistatic fitness reversal and is now beneficial.

We continue down the subgraph, checking the fitness effect of the deleterious mutation in all descendants. We first check all genotypes of equal depth, *d*, from the origin before checking those at depth *d+1* (Figure 1). If we find a descendant that does not contain the deleterious mutation we prune that descendant's children from the mutation subgraph. Eventually we will reach one of two outcomes: (1) the deleterious mutation has undergone a fitness reversal and is no longer harmful to the descendant or (2) no more descendants contain the deleterious mutation. In the second outcome, the original deleterious mutation was purged from the environment before it underwent a sign-epistatic change. In the first outcome, the original deleterious mutation became beneficial through interaction with a subsequent mutation..

Results

Deleterious Mutations contribute to long-term adaptive success in sexual populations. That occasional deleterious mutations can contribute to long-term adaptive success in asexual populations has been observed in a variety of studies (Lenski 2003, Cowperthwaite et al 2006, Covert 2010), but

[1] A unit of time in avida equal to 30 instruction executions per living organism in the population, see Ofria and Wilke (2004) for further details.

[2] Each offspring can experience at most one mutation, by design.

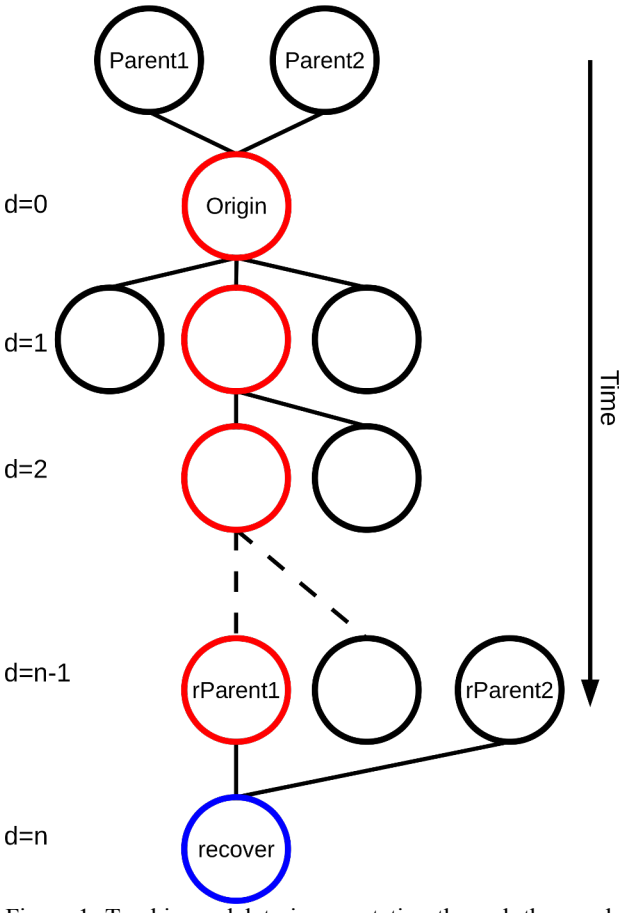

Figure 1: Tracking a deleterious mutation through the graph of descent. We mark the genotype that contains the first instance of the deleterious mutation as the "origin." The deleterious mutation that emerges in the origin genome is traced through time. At each step, the mutation's fitness effect is tested. When a fitness reversal occurs, we mark the genotype it occurs in as "recover."

not in sexual populations. To measure the impact of deleterious mutations in sexual populations, we evolved replicate sexual populations of digital organisms under two treatments, *Control* and *RvD* (revert deleterious). The Control treatment consisted of standard adaptation. The RvD treatment was identical to Control, with the exception that we monitored all mutations in offspring organisms (after division but before recombination) and determined whether an offspring organism had suffered from a deleterious (but not lethal) mutation. We reverted those offspring organisms with a deleterious mutation to the parental genotype. After reversion of a deleterious mutations, offspring organisms were subjected to recombination with other offspring organisms, as in the Control treatment.

We adapted 50 replicate populations under both treatments. We found that the dominant genotypes after adaptation had, on average, significantly higher fitness in the Control treatment than in the RvD treatment (Figure 2, p=1.74x10^{-4}, h=705.0, U-test). This finding strongly suggests that deleterious mutations contributed to the long-term evolutionary success of the Control populations. How exactly the deleterious mutations impacted adaptation remains unclear.

Classification of Individual Deleterious Mutations. How do deleterious mutations benefit long-term adaptation? We hypothesized that a fraction of deleterious mutations underwent a sign-epistatic fitness reversal. We developed an efficient algorithm to track the graph of descent in an asexual population and to test for the presence of sign-epistatic mutations on this graph. The algorithm is described in detail in the methods. In brief, the algorithm works as follows: Every mutation that lowers fitness is examined in every genotype that carries it on the graph of descent. When the mutation no longer harms its current genotype we know that it has undergone a sign-epistatic fitness-reversal. In Figure 1, the first genome that expresses the deleterious mutation is called the "origin". The "recovery" is the genotype that carries the mutation, but is the first instance where the mutation is no longer deleterious. The depth of recovery is the number of steps between the origin genotype and the recovery genotype.

Characteristics of sign-epistatic mutations in sexual populations. We ran cd our analysis of individual deleterious mutations on the first 20 replicates. Among the 6,921,517 analyzed genotypes, we found a total of 22,724 deleterious mutations. We limited our analysis to mutations that caused a fitness loss of over 1% relative to both parents, and whose effects were fully reversed via sign-epistasis. We found 902 such mutations.

Figure 3 displays the fitness cost of the 902 deleterious mutations versus their depth of recovery. Fitness cost was measured relative to the average fitness of both parents. The depth of recovery is the number of steps on the mutation subgraph from the origin of the deleterious mutation (see Figure 1). The average depth of recovery was fairly small,

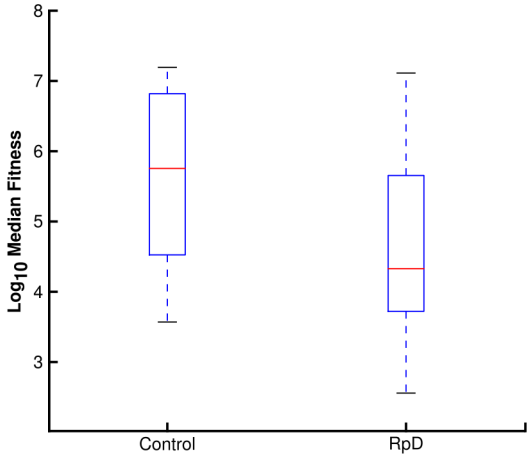

Figure 2: Effect of deleterious mutations on final organism fitness. 50 replicate populations where evolved with (control) and without (RpD) deleterious mutations. Control populations had significantly higher fitness (p=1.74x10^{-4}, h=705.0, U-test) than the RpD population. The fitness differential suggests that deleterious mutations contributed positively to adaptive evolution, most likely via fitness-reversals.

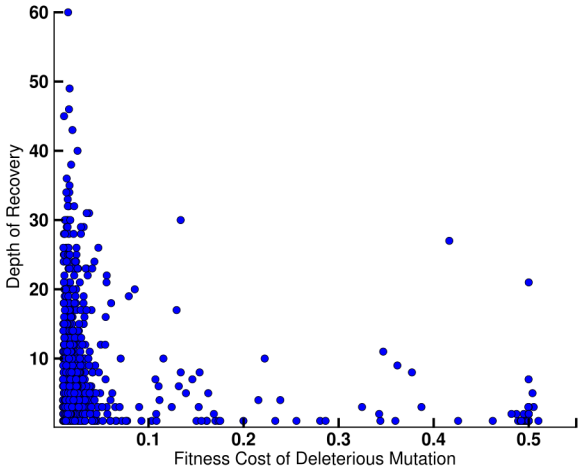

Figure 3: Fitness cost of 848 deleterious mutations (that underwent a fitness reversal) versus their depth of recovery. The x-axis shows the initial percentage of fitness loss relative to the average fitness of both parents. The y-axis shows the number of steps *d* on the graph of descent between the origin of the deleterious mutation and its recovery (see Figure 1). Most recoveries (771) occur a short distance from the origin and rescue mutations with relatively modest initial fitness cost (less than 10%).An additional 54 mutations had a fitness loss greater than 55% off fitness and are not shown.

7.80 steps on the graph of descent, although the standard deviation of depth of recovery was large, 7.93. Approximately 91.3% of recoveries took less than 20 steps, although one took 60. The average fitness cost of a deleterious mutation was 10.0% of average parent fitness, and the standard deviation was high, 22.4%. While a few deleterious mutations on the graph of descent had some extreme fitness losses (approximately 61 occurred between 50% and 99.9% fitness loss), most of the fitness losses were modest but would have been harmful had they not been compensated for.

Figure 4 shows the percent increase in fitness on recovery (i.e., the amount of fitness increase relative to the parents on the origin) versus the depth of recovery. The high increase in fitness at the time of recovery strongly suggests that the deleterious mutations contributed to the evolution of logical functions rewarded by the environment. The average increase in fitness was approximately 4.84%. The standard deviation of fitness increase was large, 434%. The smallest fitness gain was 1% while the largest fitness gain was over 3,000%. Such large fitness gains are normally only associated with the evolution of more complex logical operations (Lenski et al 2003, Covert 2010). Therefore, some deleterious mutations were likely instrumental in the evolution of complex features. The vast majority of fitness increases (747) where between 1% and 10%, suggesting optimizations of the genotypes replication efficiency.

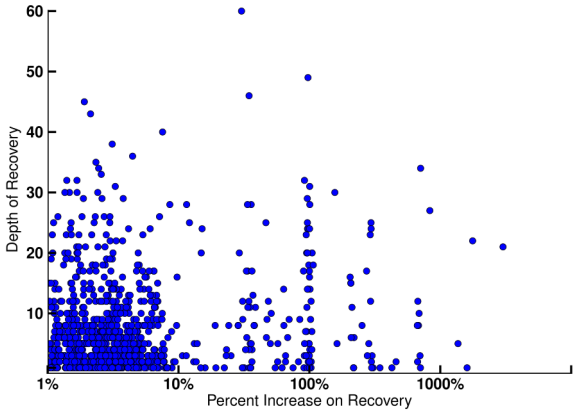

Figure 4: Percent fitness increase relative to the parent plotted against the depth of recovery. The x-axis shows the fitness increase of the recovery genotype relative to the parents of the origin (see figure 1). The y-axis shows the number of steps *d* on the subgraph between the origin of the deleterious mutation and its recovery. Most recoveries (747) are correlated with a modest but important fitness increase between 1% and 10%.

Discussion

In large asexual populations sign-epistatic mutations may appear sequentially and sweep to fixation together (Weinreich and Chao 2005, Weissman et al. 2009). Populations that experience a sweep of sign-epistatic pairs of mutations may be able to pass through fitness valleys. In sexual populations sign-epistatic mutations may be brought together by recombination, but will be disrupted by recombination unless they are tightly linked.

Analytical works have shown that recombination at low levels does not disrupt the fixation of sign-epistatic pairs, but that linkage-disequilibrium takes over at higher levels of recombination (Weinreich and Chao 2005). This implies that there is a critical recombination rate beyond which the simultenous fixation of sign-epistatic mutations is highly improbable. Our initial experiments suggest that fitness-reversals play an important role, despite a high recombination rate. While there are many other factors which must be accounted for, it seems clear that the number and frequency of sign-epistatic events is less important than where they carry the population on the fitness landscape.

We found that deleterious mutations may contribute positively to the long-term adaptation of sexual populations. Eliminating deleterious mutations from evolving populations significantly decreased the fitness of the final dominant genotypes in replicate populations. We have demonstrated that it is possible to track the entire graph of descent in sexual populations of digital organisms, from the initial ancestor to the final dominant genotype, and to track epistatic interactions among mutations on the graph. We found numerous examples of sign-epistatic recoveries in populations that experienced deleterious mutations. In these examples, mutations caused a fitness loss in the genetic backgrounds in which they arose,

but interacted epistatically with subsequent mutations and eventually contributed positively to fitness.

Our sexual populations were strictly haploid, the introduction of diploidy or polyploidy could create the added complexity of considering dominant and recessive traits. In addition, we had only a single crossover point. The effect we observed may be disrupted as additional crossover points increase linkage-disequilibrium between traits. However, previous work in digital organisms has shown that recombination encourages organisms to evolve more modular genotypes (Misevic et al. 2006). Therefore, additional crossover points may also encourage mutations effecting a single trait to be more tightly linked on the genome. Further work is necessary to resolve this issue.

Our work indicates that there are circumstances when sign-epistasis may play an important role in the evolution of sexual populations. It opens new possibilities for researching the exploration of fitness landscapes in sexual populations. Asexual studies of fitness landscapes generally assume that populations will move from one genotype to a genotype separated by only a few point mutations. In sexual systems, populations may move in great leaps and bounds across the fitness landscape, due to recombination. Our graph of descent gives us the ability to observe, for the first time, the movement of sexual populations through fitness landscapes.

Acknowledgments

This material is based in part upon work supported by the National Science Foundation under Cooperative Agreement No. DBI-0939454 and by The Freshman Research Initiative at the University of Texas at Austin. The authors acknowledge Charles Ofria and Richard Lenski for valuable conversations.

References

Covert III, A. W. (2010), *On the Beneficial Effects of Deleterious Mutations.* Ph.D thesis, Department of Computer Science Michigan State University.

Cowperthwaite MC, Bull JJ, Meyers LA (2006), From bad to good: Fitness reversals and the ascent of deleterious mutations. *PLoS Comp Biol* 2:1292-1300.

Iwasa Y, Michor F, Nowak MA (2004) Stochastic tunnels in evolutionary dynamics. *Genetics* 166:1571-1579.

Kvitek DJ, Sherlock G (2011), Reciprocal Sign Epistasis between Frequently Experimentally Evolved Adaptive Mutations Causes a Rugged Fitness Landscape. *PLoS Genet* 7:e1002056.

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

Lenski RE, Ofria C, Pennock RT, Adami C (2003), The evolutionary origin of complex features. *Nature* 423:139-144.

Misevic D, Ofria CA, Lenski RE (2006) Sexual reproduction reshapes the genetic architecture of digital organisms. *Proc R Soc London B* 273:457-464.

Ofria C, Wilke CO (2004), Avida: a software platform for research in computational evolutionary biology. *J Artif Life* 10:191-229.

van Nimwegen E, Crutchfield JP (2000), Metastable evolutionary dynamics: Crossing fitness barriers or escaping via neutral paths? *Bull Math Biol* 62:799-848.

Weinreich DM, Chao L (2005), Rapid evolutionary escape by large populations from local fitness peaks is likely in nature. *Evolution* 59:1175-1182.

Weissman DB, Desai MM, Fisher DS, Feldman MW (2009) *J Theor Biol* 75:286-300.