doi:10.1111/j.1558-5646.2007.00064.x

EFFECTS OF POPULATION SIZE AND MUTATION RATE ON THE EVOLUTION OF MUTATIONAL ROBUSTNESS

Santiago F. Elena,[1,2] Claus O. Wilke,[3,4] Charles Ofria,[5,6] and Richard E. Lenski[7,8]

[1]*Instituto de Biología Molecular y Celular de Plantas, CSIC-UPV, 46022 Valencia, Spain*

[2]*E-mail: sfelena@ibmcp.upv.es*

[3]*Section for Integrative Biology and Center for Computational Biology and Bioinformatics, University of Texas, Austin, Texas 78712*

[4]*E-mail: cwilke@mail.utexas.edu*

[5]*Department of Computer Sciences and Engineering, Michigan State University, East Lansing, Michigan 48824*

[6]*E-mail: ofria@msu.edu*

[7]*Department of Microbiology and Molecular Genetics, Michigan State University, East Lansing, Michigan 48824*

[8]*E-mail: lenski@msu.edu*

Received October 17, 2006
Accepted November 16, 2006

It is often assumed that the efficiency of selection for mutational robustness would be proportional to mutation rate and population size, thus being inefficient in small populations. However, Krakauer and Plotkin (2002) hypothesized that selection in small populations would favor robustness mechanisms, such as redundancy, that mask the effect of deleterious mutations. In large populations, by contrast, selection is more effective at removing deleterious mutants and fitness would be improved by eliminating mechanisms that mask the effect of deleterious mutations and thus impede their removal. Here, we test whether these predictions are supported in experiments with evolving populations of digital organisms. Digital organisms are self-replicating programs that inhabit a virtual world inside a computer. Like their organic counterparts, digital organisms mutate, compete, evolve, and adapt by natural selection to their environment. In this study, 160 populations evolved at different combinations of mutation rate and population size. After 10^4 generations, we measured the mutational robustness of the most abundant genotype in each population. Mutational robustness tended to increase with mutation rate and to decline with population size, although the dependence with population size was in part mediated by a negative relationship between fitness and robustness. These results are independent of whether genomes were constrained to their original length or allowed to change in size.

KEY WORDS: **Deleterious mutations, digital organisms, fitness landscapes, mutational robustness, neutral networks, population size.**

Mutation is a double-edged sword. On one side, it is the ultimate source of genetic variation and the raw material for adaptation by natural selection; a lineage without any mutations would be fated to extinction because of its inability to respond to environmental change. On the other side, mutations typically reduce an organism's fitness, and most are eliminated from populations by selection against them. Because mutation is a random process, organisms cannot benefit from mutations without also suffering their negative effects. However, mechanisms that minimize the detrimental effects of mutations may be selectively advantageous. In

other words, organisms may benefit from mechanisms that promote robustness by buffering against the pernicious effects of deleterious mutations, especially at high mutation rates (de Visser et al. 2003).

Krakauer and Plotkin (2002) hypothesized that robustness can be achieved by two distinct mechanisms, each operating over a different range of population sizes. Organisms that exist as small populations, such as those typical of many multicellular organisms, evolve robustness by masking the harmful effects of deleterious mutations, often by making certain functions redundant. Examples of such mechanisms include gene duplication, alternative metabolic pathways, or chaperone proteins that buffer against mutation-induced problems in other enzymes. All these mechanisms would produce phenotypes that are similar or identical to the unmutated wild-type, such that affected individuals would have similar chances for survival and reproductive success. By contrast, lineages that exist as very large populations, such as most viruses and bacteria, may become more robust simply by avoiding the costs of these redundancy mechanisms. Although this strategy would lead to individual-level hypersensitivity to deleterious mutations, those individuals carrying harmful mutations would be readily eliminated from large populations while preserving unmutated genomes, which are more fit. Krakauer and Plotkin (2002) referred to such hypersensitivity as antiredundancy, and possible examples include overlapping reading frames, haploidy, and the loss of systems for genome repair. In the rest of this paper, we will use the terms robustness and nonrobustness instead of redundancy and antiredundancy, respectively, as used by Krakauer and Plotkin (2002). We think that our wording is more appropriate because some of the mechanisms described above do not necessarily imply multiple copies of the same gene function, as suggested by the term redundancy.

The Haldane–Muller principle states that the average equilibrium fitness of an infinite population of haploid asexual organisms is given by $e^{-\mu}$, where μ is the mutation rate, and this equilibrium is independent of the precise geometry of the fitness landscape (Bürger 2000). However, if neutral mutations are considered, then this approximation is no longer valid, and the average equilibrium fitness depends on the geometry of the fitness landscape (Wilke and Adami 2003). In that case, another selective pressure comes into play, especially at high mutation rates, that pushes an evolving population toward regions of the landscape where the density of neutral mutations is higher (van Nimwegen et al. 1999; Wilke 2001a) and the average genome is more robust with respect to the effects of new mutations. This phenomenon of selection for robustness at high mutation rates can be understood in two equivalent ways (de Visser et al. 2003; Edlund and Adami 2004). On one hand, it can be understood as a pressure for populations of evolving genomes to occupy more highly connected areas of a neutral network. On the other hand, within a more general fitness-landscape paradigm, it can be viewed as a pressure to occupy broad rather than narrow fitness peaks; genomes that occupy narrow fitness peaks are easily mutated to low fitness, whereas genomes that are on broad peaks are less susceptible to deleterious mutations. This latter view gives rise to the notion of the "survival of the flattest," which has been postulated theoretically (Schuster and Swetina 1988; Wilke 2001a) as well as shown to operate in both digital organisms (Wilke et al. 2001) and simulated RNA evolution (Wilke 2001b). Indeed, we would suggest that low-flat and high-steep adaptive peaks correspond to the alternative strategies of redundancy and antiredundancy proposed by Krakauer and Plotkin (2002).

Here, we investigate evolving populations of digital organisms to examine how population size and mutation rate, as well as restrictions to changes in genome length, affect the evolution of genomic robustness against deleterious mutations. The use of digital organisms to address questions in evolutionary biology began only recently (Ray 1991; Maynard Smith 1992; Adami 1998, 2006; Lenski et al. 1999, 2003; Lenski 2001; Wilke et al. 2001; Yedid and Bell 2001, 2002; Wilke and Adami 2002; O'Neill 2003; Chow et al. 2004; Misevic et al. 2004, 2005; Ostrowski et al. 2006), but has several potential advantages. First, digital organisms allow one to perform experiments on various scales that are beyond reach with any biological entity, including microbes. Second, one can perform certain manipulations on digital organisms that are physically impossible to do with natural organisms. Third, one can collect much more detailed and precise information about an evolving digital system than a natural one. Fourth, these populations are real evolving systems, not mere simulations; digital organisms truly evolve new and sometimes surprising functionalities. Fifth, experiments with digital organisms allow one to test evolutionary concepts and hypotheses that extend beyond the organic forms with which biologists are familiar. Finally, it is important to stress that the Avida implementation of Darwinian processes radically differs from genetic algorithms (or its variations such as genetic programming and evolutionary strategies) in three main characteristics (Ofria et al. 2002; Wilke and Adami 2002; Ofria and Wilke 2004; Adami 2006). First, there is no explicit fitness function for the digital organisms; fitness is measured as the rate at which an organism produces offspring, and the expected value of that rate never feeds back into the experimental system. Second, the genomes of the digital organisms are computer programs that have the potential, at least in principle, to perform any possible computation. Hence, there is a nontrivial genotype–phenotype map, allowing for open-ended evolution of complex traits. Finally, organisms in Avida directly interact with one another, although the current experiments limit these interactions to scramble competition for simplicity of analysis.

Material and Methods

THE AVIDA PLATFORM

The Avida platform is software in which self-replicating computer programs—that is, digital organisms—evolve by means of the basic population-genetic forces of mutation, selection, and random drift (Adami 1998; Ofria and Wilke 2004). To replicate, digital organisms must copy their genomes line by line, and then divide the resulting genomes to yield two newly independent organisms. As with biochemical organisms, replication is subject to errors including point mutations as well as insertions and deletions. For point mutations, a random instruction is written in place of the original one, whereas for insertion or deletion mutations, a random instruction is either added to or removed from the genome.

The success of a digital organism depends on its realized growth rate relative to other organisms in its population. An organism's growth rate in a given environment depends on the rate at which it garners energy from resources and its efficiency in converting that energy into progeny. The Avida environment for any given experiment is set by the investigator and determines the mutation rates as well as the collection of resources that an organism can use to gain energy and increase its replication rate. Resources are converted to energy by the performance of mathematical computations (Wilke and Adami 2002; Lenski et al. 2003). By performing such computations, also termed *functions*, organisms gain energy (extra CPU cycles) that they can use to accelerate their replication. Performing functions and gaining CPU cycles are therefore equivalent to metabolic reactions that yield energy. Resources are equally available to all organisms, similar to the situation in a well-stirred chemostat. In all experiments reported here, the resources associated with each function were unlimited, so that all interactions reflect scramble competition for CPU cycles, which are the sole limiting resource. With this setup, the Avida world has only a single ecological niche, and stable coexistence based on partitioning different resources is therefore precluded. Population size was constant in each run; this constancy was maintained because the placement of each new offspring in a population caused the death of some randomly chosen individual. No particular digital organism is explicitly favored in any experiment; each organism's phenotype depends on the complex rules that govern how a genomic program is executed, and selection depends on the interaction between the resulting organismal phenotypes and the environment.

The experiments reported here were performed using version 1.99 of the Avida software. This software is freely available from devolab.cse.msu.edu/software/avida/. In any given experimental treatment, each replicate evolving population had identical initial conditions except for the random-number seed. This initial seed causes runs to differ at all subsequent points where stochastic events occur, including mutations and the physical location where each new offspring is placed in the population.

The genome of the ancestral organism (designated *organism.heads.100*) used in all our experiments was 100 instructions in length. This organism could self-replicate but was unable to perform any logical computations. Specific details about the Avida software, including the default set of instructions and the definition of the environment, can be found in Ofria and Wilke (2004) and at the website shown above. Experiments ran under the Linux operating system on either a Beowulf cluster of 96 AMD Athlon 1600+ processors (at MSU) or an IBM 1350 cluster of 120 Intel Xeon processors (at UPV).

EXPERIMENTAL TREATMENTS

Each treatment combined different values of population size and mutation rate to examine their effects on the evolution of mutational robustness. Also, every experiment was run under two different conditions regarding genome length L; in the standard runs L could change during evolution, whereas in the runs with genome-length restriction (henceforth GLR) length was fixed at the ancestral value of 100 instructions. The treatments had two point mutation rates, μ, equal to 0.002 or 0.02 per instruction per generation, and four population sizes, N, of 10^2, 10^3, 10^4, or 5×10^4 individuals (with the upper end set by computational limitations). In those experiments where L was allowed to change, the rates of insertion and deletion mutations were set to 0.05 per genome per generation; in these experiments the total genomic point mutation rate, $U = \mu L$, therefore also evolved. By contrast, in the GLR experiments, U was either 0.2 or 2 throughout any run.

We ran 10 replicate populations for each combination of N, μ, and GLR, giving 160 experimental populations in all. Each population evolved for 10^5 updates, where an update is an arbitrary unit of time in which $30 \times N$ instructions are executed globally (and thus organisms execute an average of 30 instructions each). A typical generation requires about four to 16 updates, depending on an organism's genome size, phenotypic complexity, and efficiency of execution. A zip file containing all the Avida configuration and results files can be downloaded from bioinfo.ibmcp.upv.es/downloads/runs.tar.gz.

MEASURING THE FITNESS OF DIGITAL ORGANISMS

We define the fitness, W, of a genotype in a particular environment by its expected rate of offspring production. Fitness was therefore measured by dividing the rate at which an organism acquires the energy used to execute the instructions in its genome by the number of instructions that it must execute to produce an offspring (Adami 1998). This energy is obtained as discrete quanta called single-instruction processing units, or SIPs (Lenski et al. 2003;

Adami 2006). The rate at which an organism acquires energy is proportional to its genome length and a multiplier based on interactions between the organism's computational metabolism and its environment; in particular, an organism receives more SIPs if it performs certain computational functions in an environment that rewards those functions (Lenski et al. 2003). This definition allows us to compare the relative rates of reproduction of any two organisms. Thus, if one genotype has W twice as large as another genotype, then the former is expected to produce offspring at twice the rate of the latter. Genotypes that cannot produce any offspring, owing to certain mutations, have zero fitness.

The absolute fitness of the ancestral organism, *organism.heads.100*, is 0.2494. This value was used to normalize the fitness of the evolved populations relative to the ancestor; for example, a relative fitness of 1.5 means that an evolved organism produces offspring at a rate 50% faster than did the ancestor.

QUANTIFYING MUTATIONAL ROBUSTNESS

We quantified evolutionary changes in mutational robustness as follows. At the end of the experiment, we identified the most abundant genotype from each independently evolved population. We then systematically constructed and tested the fitness effects of every possible single mutation in each of these 160 genetic backgrounds, as previously described (Lenski et al. 1999). The fitness of each mutant relative to its nonmutant parent was expressed as $s_i = W_i/W_0 - 1$, where W_i is the fitness estimated for the ith mutation and W_0 is the fitness of the parent. The average selection coefficient, $\bar{s}$, was calculated over all single mutations. Mutational robustness, as generally discussed, concerns neutral, deleterious, and lethal mutations only. Hence, we discarded those rare mutations with beneficial effects (i.e., $s_i > 0$) from the calculations of $\bar{s}$ reported in this paper. (Additional analyses indicate, however, that including beneficial mutations does not qualitatively affect our conclusions [data not shown].) If a population has evolved mutational robustness, then mutations should, on average, have a little or no impact on fitness and; thus, $\bar{s}$ would be small. By contrast, the less robust a genotype is, the larger the average effect across all possible mutations would be. Therefore, the quantity $\bar{s}$ constitutes our measure for mutational sensitivity, and $1/\bar{s}$ can be viewed as a measure for robustness.

A second test was carried out for those populations that showed the lowest and the highest average sensitivity by the above criterion. The most abundant genotype from each of these populations was allowed to evolve for an additional 10^3 updates under modified conditions where only deleterious mutations were allowed to occur. To achieve such conditions, we eliminated the rewards for performing any computational functions that had not yet been evolved by the organisms, so that any mutations that yielded those new functions would no longer be beneficial. Also, population size was reduced to 100 regardless of its previous size, such that genetic drift became a potent evolutionary force. These two changes ensured that organisms would not acquire any new beneficial functions, but they could accumulate deleterious mutations that impacted their existing functions. Fitness values were measured at the start and at the end of the second experiment. If organisms had not evolved mutational robustness during the first experiment, then their fitness levels should decline sharply during the second experiment. By contrast, if organisms evolved robustness during the first experiment, then they should be buffered against the harmful effects of new mutations and their fitness should decline to a lesser extent.

STATISTICAL ANALYSES

To examine the effects of N, μ, and GLR on W and $\bar{s}$, we fit a three-way factorial analysis of covariance (ANCOVA) to the data. Both μ and GLR were treated as fixed factorial factors, whereas $\log N$ was treated as a covariate. We were interested in testing the three main effects, the interaction between the two factorial factors, and all possible differences in slopes (first- and second-degree interactions with the covariable). Type III sums of squares were used. For every treatment, each one of the 10 independent replicate populations was treated as a replicate. All statistical analyses were performed using SPSS version 14 software.

Results

EFFECTS OF POPULATION SIZE, MUTATION RATE, AND GENOME LENGTH ON FITNESS EVOLUTION

Figure 1 shows the effects of population size N, mutation rate μ, and GLR on the fitness W of the most abundant genotypes at the end of the first evolution experiment. Notice that both N and W are shown with values decimal log-transformed. Table 1 shows the corresponding ANCOVA using the log-transformed W data. The main results of this experiment can be summarized as follows. First, N had a highly significant effect on the W that was achieved (Table 1, $P < 0.0001$), with the most abundant genotypes from $N = 5 \times 10^4$ populations reaching, on average, twice the fitness of those from $N = 100$ populations (Fig. 1).

Second, μ also had a highly significant effect on W (Table 1, $P < 0.0001$), with the higher mutation rate producing higher fitness levels at the end of the experiment (Fig. 1). Genomes that evolved at the high $\mu = 0.02$ were, on average, 23.50% more fit than those from the same size populations that evolved at the $\mu = 0.002$. The positive effect of producing more beneficial mutations was evidently greater than the negative effect of generating more deleterious mutations over the mutation rates and other conditions of our experiment. Furthermore, the fitness effects of μ and N were not independent, as indicated by a significant $\mu \times \log N$ interaction (Table 1, $P = 0.0457$). Genomes evolved at the high mutation rate were 61.58% times fitter than those evolved at the

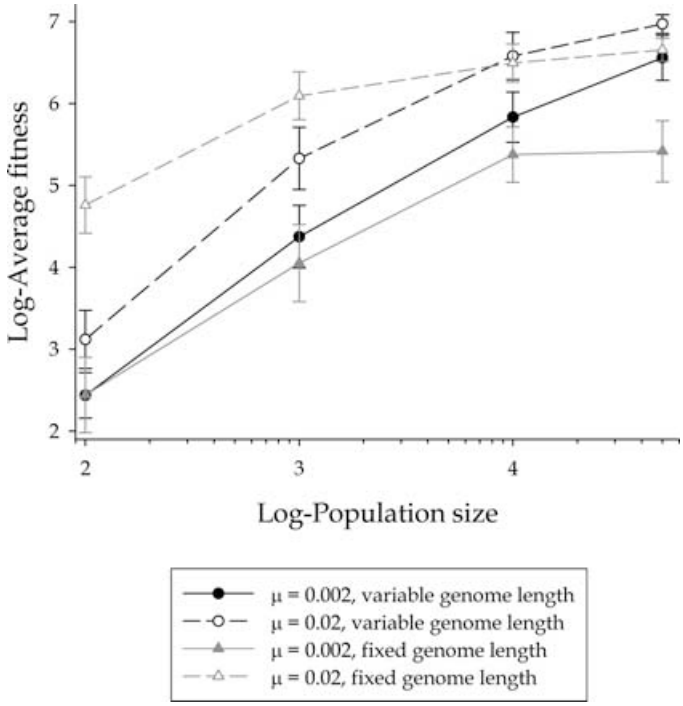

Figure 1. Effects of population size, mutation rate, and genome-length restriction on log-transformed fitness of the most abundant genotypes isolated at the end of the evolution experiments. Error bars represent standard errors of the mean ($n = 10$).

low mutation rate when $N = 100$, but the corresponding difference was only 13.77% when $N = 5 \times 10^4$.

Third, the effects of GLR on fitness evolution were rather more complex. The overall main effect of GLR was small (2.07%) although significant (Table 1, $P = 0.0023$). The interaction between GLR and log N was also significant; not restricting genome length allowed 12.06% faster evolution in the largest populations but not in smaller ones (Fig. 1). There was another significant interaction between μ and GLR (Table 1, $P = 0.0460$), with populations evolved at high mutation rate being, on average, 9.12% fitter if genome length was held constant than if it was allowed to change. The three-way interaction among μ, GLR, and log N was not significant (Table 1, $P = 0.1602$).

Table 1. Analysis of covariance of the log-transformed fitness data shown in Figure 1. Mutation rate (μ) and genome-length restriction (GLR) were treated as fixed factors; log-population size (log N) was treated as a covariate. Type III sums of squares were used. Overall $R^2 = 0.6234$ including significant effects only.

Source	SS	df	F	P
log N	1327.0559	1	213.7086	<0.0001
μ	87.6338	1	14.1125	0.0002
μ×log N	25.2064	1	4.0592	0.0457
GLR	59.8095	1	9.6317	0.0023
GLR×log N	57.8781	1	9.3207	0.0027
μ×GLR	25.1261	1	4.0463	0.0460
μ×GLR×log N	12.3659	1	1.9914	0.1602
Error	943.8670	152		

At large N ($\geq 10^4$), the final log-transformed fitness shows evidence of diminishing returns with increasing N at both low and high μ, and irrespective of GLR. In fact, a hyperbolic model (which levels off at some maximum value) fits the data significantly better than a simple linear model, despite using an extra degree of freedom (partial F-tests: in all four cases $F_{1,37} \geq 20.1502$, $P < 0.0001$).

MUTATIONAL SENSITIVITY EVOLVES IN LARGE POPULATIONS, WHEREAS MUTATIONAL ROBUSTNESS EVOLVES AT HIGH MUTATION RATES

Over the parameter ranges used in our experiments, both larger N and higher μ had positive effects on fitness evolution. By contrast, as we will now show, N and μ exerted opposing effects on the evolution of genomic robustness to mutations. Our first test for mutational robustness examined the sensitivity of genotypes evolved under the various treatments to all possible single mutations, as described in Material and Methods. Figure 2 shows the mutational sensitivity, $\bar{s}$, measured for the most abundant genotypes at the end of the experimental evolution as a function of treatment variables N, μ, and GLR. Table 2 shows the results of the corresponding ANCOVA. Several conclusions can be drawn from this analysis.

First, log N had a highly significant effect on the evolution of mutational sensitivity (Table 2, $P < 0.0001$), with larger populations evolving greater sensitivity (lower robustness) to mutations

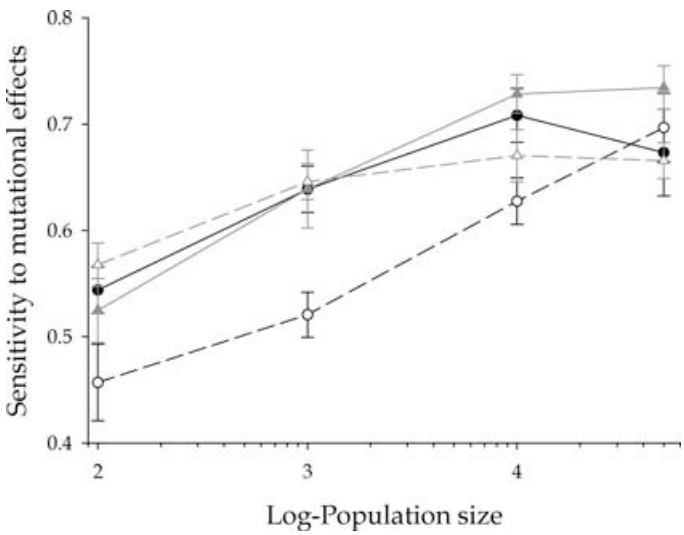

Figure 2. Effects of population size, mutation rate, and genome-length on the mutational sensitivity of the most abundant genotypes isolated at the end of the evolution experiments. Symbols and lines are as described in Figure 1.

Table 2. Analysis of covariance of the mutational sensitivity data shown in Figure 2. See Table 1 for further details. Overall $R^2 = 0.4381$ including significant effects only.

Source	SS	df	*F*	*P*
log *N*	0.7034	1	99.8090	<0.0001
μ	0.0032	1	0.4473	0.5046
μ×log *N*	0.0005	1	0.0643	0.8002
GLR	0.0228	1	3.2299	0.0743
GLR×log *N*	0.0071	1	1.0124	0.3159
μ×GLR	0.0876	1	12.4303	0.0006
μ×GLR×log *N*	0.0703	1	9.9780	0.0019
Error	1.0713	152		

(Fig. 2). The average sensitivity for genomes evolved at $N = 5 \times 10^4$ was 32.34% greater than that observed for genomes evolved at $N = 100$ (Fig. 2).

Second, and surprisingly, μ had no direct effect on mutational sensitivity (Table 2, $P = 0.5046$), nor was its interaction with log *N* significant (Table 2, $P = 0.8002$). However, more complex effects of μ on mutational robustness will be identified in the next paragraph and in the following section.

Third, as in the case of fitness, the effects of GLR on sensitivity were complex. First, GLR had no overall effect on evolved mutational sensitivity (Table 2, $P = 0.0743$) nor did GLR significantly interact with population size (Table 2: GLR × log *N*, $P = 0.3159$). However GLR did show a significant interaction with mutation rate (Table 2: μ × GLR, $P = 0.0006$). Furthermore, the magnitude of this interaction was affected by *N*, as shown by a significant three-way interaction (Table 2: μ × GLR × log *N*, $P = 0.0019$). In populations that evolved at the lower μ, restricting genomes to a fixed length led to a 5.87% increase in mutational sensitivity (reduced robustness) only when $N > 100$ (Fig. 2, solid lines). At the higher μ, however, GLR had a larger negative effect when $N \leq 10^4$, leading to 17.36% increase in sensitivity, but it had the opposite effect (15.67% more robust) in the largest *N* (Fig. 2, dashed lines). Thus, the effects of GLR, μ, and *N* were interdependent and complex.

Across all 16 treatments, the most sensitive genomes (highest $\bar{s}$) evolved at $N = 5 \times 10^4$ and $\mu = 0.002$ with genome length fixed, whereas the most robust genomes (lowest $\bar{s}$) evolved at $N = 100$ and $\mu = 0.02$ with varying genome length. The dominant evolved genotypes from these 20 populations were employed in the second test of mutational robustness, which used a mutation-accumulation experiment to prevent adaptive evolution as described in Material and Methods. At the end of the mutation-accumulation phase, the average log *W* for the most sensitive genotypes had declined by about 9.69% relative to their immediate progenitors, and this loss was highly significant (paired *t*-test: $t_9 = 8.1871$, $P < 0.0001$). By contrast, the most robust populations showed no significant loss of log *W* during the mutation-accumulation experiment relative to their progenitors (paired *t*-test: $t_9 = 1.2296$, $P = 0.2500$). This second test therefore supports the premise of the first set of experiments that robustness can be quantified by examining the fitness effects of all one-step mutational neighbors. It also indicates that the combination of high μ, small *N*, and variable length favored the evolution of robust genomes, whereas the combination of low μ, large *N*, and fixed length led to genomes that were more sensitive to the harmful effects of random mutations.

MUTATIONAL SENSITIVITY IS CORRELATED WITH FITNESS INCREASE

Those treatments that allowed populations to evolve higher fitness also produced genomes that were more sensitive to the harmful effects of random mutations, as seen by the overall similarity in form in Figures 1 and 2. However, it is not clear whether this similarity reflects parallel, but largely independent, treatment effects on these two dependent variables; alternatively, the similarity might suggest some deeper, and presumably mechanistic, coupling between higher fitness and greater sensitivity. To explore this issue further, we computed the partial correlation coefficient between log-transformed fitness and mutational sensitivity while controlling for the three variables that were manipulated in the evolution experiment (*N*, μ, and GLR) (Sokal and Rohlf 1995). The resulting correlation was highly significant ($r = 0.6081$, 155 df, $P < 0.0001$). This finding confirms that the more fit genotypes also tended to become more sensitive (less robust) to the harmful effects of random mutations, even among the replicate populations that experienced the same treatment during their evolution.

One possible reason for this correlation could be that more fit organisms are inherently more prone to damage by mutation. If this were the case, then the observed increases in mutational sensitivity with *N* would not be a direct consequence of selection to minimize costs of redundancy and other mechanisms to achieve robustness, but rather these trends would reflect a correlated response to the selective advantage of encoding more computational functions in the genome. A statistical way to eliminate the effect of fitness on mutational sensitivity, while testing for the effects of μ, *N*, and GLR, is to incorporate log *W* as a covariate in the analysis of covariance of the sensitivity data. If the results shown in Table 2 were only a consequence of the correlation between log *W* and $\bar{s}$, then the significant effects of the other factors would disappear. Table 3 presents this new analysis, whereas Figure 3 shows the magnitude of the effects of the other factors on mutational sensitivity after removing the effect of the covariate log *W*. In agreement with the previous correlation, log *W* significantly affected the magnitude of $\bar{s}$. Interestingly, after removing the effect of log *W* on $\bar{s}$, the main effect of log *N* remained significant (Table 3, $P = 0.0065$). Genomes evolved in small populations ($N = 100$)

Table 3. Analysis of covariance of the mutational sensitivity data shown in Figure 2, but incorporating log *W* as a covariate to account for the observed correlation between fitness and mutational sensitivity. See Table 1 for further details. Overall $R^2 = 0.2870$ including significant effects only.

Source	SS	df	*F*	*P*
log *W*	0.1858	1	31.6909	<0.0001
log *N*	0.0446	1	7.6053	0.0065
μ	0.0322	1	5.4860	0.0205
μ×log *N*	0.0024	1	0.4014	0.5273
GLR	0.0017	1	0.2878	0.5924
GLR×log *N*	0.0005	1	0.0797	0.7780
μ×GLR	0.0496	1	8.4582	0.0042
μ×GLR×log *N*	0.0460	1	7.8423	0.0058
Error	0.8855	151		

were, on average, 32.34% less sensitive to mutations than those evolved in large populations ($N = 5 \times 10^4$), independent of the fitness gains achieved for each population size (Fig. 3). Moreover, the main effect of μ becomes significant (Table 3, $P = 0.0205$), with populations evolved at high mutation rate being, on average, 6.53% less sensitive to mutational effects than those evolved at low mutation rate (Fig. 3). The interaction μ × GLR (Table 3, $P = 0.0042$) and the three-way μ × GLR × log *N* interaction (Table 3, $P = 0.0058$) also remained significant, confirming the importance of genome-length variation for also influencing the evolution of mutational robustness (Fig. 3).

To summarize this section, small populations evolved greater robustness to mutations than large ones. This effect was partly, but not entirely, a correlated response that reflected the higher fitness achieved in large populations and the resulting greater susceptibility of those high-functionality genomes to deleterious mutations. Population size also impacted the evolution of robustness more directly (i.e., after controlling for fitness as a covariate), as did mutation rate and interactions with GLR.

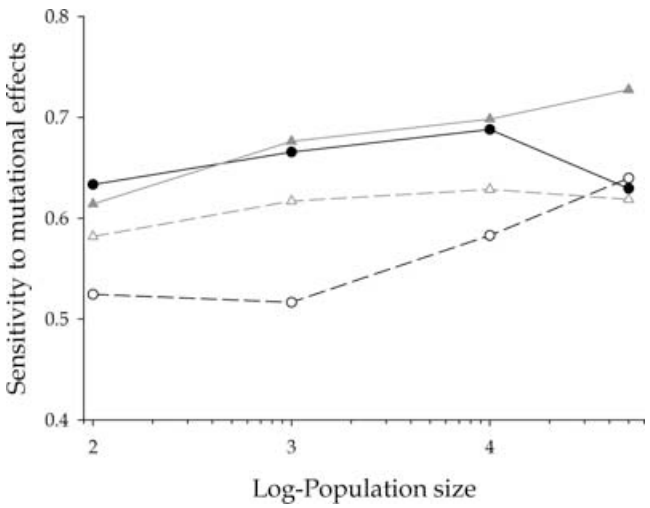

Figure 3. Effects of population size, mutation rate, and genome-length on mutational sensitivity after removing the effect of the covariate log *W* (Table 3). Symbols and lines are as described in Figure 1.

Discussion

Genomes of different species vary in their robustness to mutations. Understanding the complex interplay of evolutionary forces that shape mutational robustness has emerged as an interesting, diverse, and complex area of research (Kondrashov 1982; Scharloo 1991; Stearns and Kawecki 1994; Barkai and Leibler 1997; Elena and Lenski 1997, 2001; Nowak et al. 1997; Wagner et al. 1997; Lenski et al. 1999; van Nimwegen et al. 1999; Wagner 1999, 2005; Wagner and Stadler 1999; Gibson and Wagner 2000; Kawecki 2000; Wilke 2001b; Wilke and Adami 2001 2003; Wilke et al. 2001; Krakauer and Plotkin 2002; de Visser et al. 2003; Edlund and Adami 2004; Hermisson and Wagner 2004; Misevic et al. 2005). In the paper that stimulated our experiments, Krakauer and Plotkin (2002) argued for important effects of population size on the evolution of mutational robustness. In smaller populations, such as typical of most eukaryotes, they suggested that selection would favor functional redundancy as a means to promote robustness. In that case, deleterious mutations can persist in a population without any apparent phenotypic effect. By contrast, in larger populations, such as those of microorganisms, selection would minimize redundancy and other such mechanisms to avoid the associated costs. The result would be high sensitivity to mutational effects, making it easier for selection to eliminate deleterious mutations.

In this study, we performed experiments using digital organisms to investigate the interplay between mutation rate and population size in the evolution of mutational robustness. The direct selective advantage of robustness is expected to be of the same order of magnitude as the mutation rate (Wagner et al. 1997; Proulx and Phillips 2005). Consistent with that expectation, digital genomes that evolved at low mutation rates were more sensitive (less robust) to mutational effects than those that experienced high mutation rates during the evolution experiment (Figs. 2 and 3). Indeed, for robustness to evolve as an adaptive response to recurrent deleterious mutations, population size must be larger than the inverse of the mutation rate (i.e., $N\mu \gg 1$), and this condition makes the evolution of robustness difficult to achieve especially in small populations (Wagner 2005; and references therein). However, robustness might evolve not site by site but by more global mechanisms, such as redundancy, that might hide mutational effects throughout an entire genome (Krakauer and Plotkin 2002). In that case, robustness could be achieved more readily, and would be especially important in small populations wherein selection is less

efficient at purging deleterious mutations. Indeed, digital genomes evolved in small populations were more robust to deleterious mutations than those evolved in large populations (Figs. 2 and 3, Tables 2 and 3). Thus, genomic robustness of digital organisms tended to increase directly with mutation rate and inversely with population size, consistent with the predictions of Krakauer and Plotkin (2002) as well as more general theoretical expectations.

An additional and important complication, however, arises from the relationship between fitness and mutational robustness. The same variables that were predicted to affect the evolution of robustness also influenced the evolution of organismal functions that contribute to fitness. In particular, both large population size and high mutation rate promoted the evolution of computational functions that contributed to high fitness while also increasing susceptibility to mutations, because high-fitness genomes have more targets available for deleterious mutations. We addressed this complication by performing analyses in which log-transformed fitness served as a covariable to identify those factors that influenced robustness independently of their fitness effects (Table 3 and Fig. 3). The addition of this covariable confirmed and clarified the overall effects of population size and mutation rate on the evolution of robustness, although interactions between variables remained as added complications.

GLR, in particular, was involved in significant interactions with both mutation rate and population size. Even so, the trends of greater mutational robustness at high mutation rate and in small populations were observed whether genome size was allowed to evolve or, alternatively, was restricted to the ancestral length. Therefore, it is not necessary to invoke changing genome size to explain those trends. However, when genome length was allowed to evolve freely, the resulting effect was again qualitatively consistent with the theory of Krakauer and Plotkin (2002). That is, their theory predicts that small populations should evolve longer, more redundant genomes than for large populations. The pertinent test in our study applies to those populations that evolved at high mutation rate, because mutation pressure provides the selection for robustness and redundancy. The ancestral genome was 100 instructions in length. The average length of the evolved organisms in the smallest populations ($N = 100$) increased to 121.9 ($\pm$6.5 SEM), whereas the average for the largest populations ($N = 5 \times 10^4$) was only 95.9 ($\pm$7.8 SEM), and this difference is significant in the predicted direction ($t_{18} = 2.5553$, one-tailed $P = 0.0099$). But having a longer genome is a double-edged sword: although it promotes redundancy and thus mutational robustness, it also imposes slower genome replication. It is not surprising that larger populations generally evolved higher fitness than did smaller populations (Fig. 2). In addition to the increased number of beneficial mutations in large populations, and the greater efficacy of selection in removing deleterious mutations in large populations (Kimura 1962), our experiments indicate that large population size evolved more streamlined genomes that replicate faster. This outcome, too, supports the hypothesis of Krakauer and Plotkin (2002) that large populations should avoid the costs of redundancy mechanisms.

Selection for genome robustness can drive evolving populations toward regions in sequence space that are relatively flat, so that more mutations are neutral than in other regions (van Nimwegen et al. 1999; Wilke 2001a; Wilke et al. 2001). In agreement with this prediction, the average fraction of neutral mutations in the most robust populations ($N = 100$, high μ, and without GLR) was 36.45%, whereas it was only 12.72% in the most sensitive ones ($N = 5 \times 10^4$, low μ, and with GLR), and this difference is highly significant ($t_{18} = 8.8940$, $P < 0.0001$). We have also shown that selection for robustness is especially strong in small populations that experience high mutation rates. These patterns may contribute to differences among organisms in nature in the frequency of neutral sites in their genomes (Ofria et al. 2003). These differences may, in turn, influence evolvability (Wagner and Altenberg 1996; Earl and Deem 2004) because genomes with more neutral sites should drift more easily along neutral networks and might thereby evolve new functions more readily (Huynen et al. 1996; Wilke 2001a; but for a different view, see Lenski et al. 2006).

Despite recent advances in genomics, transcriptomics, and proteomics, it remains very difficult to investigate the evolutionary forces that shape such important genomic features as modularity and robustness. As our study and previous work show, digital organisms provide an excellent alternative to natural systems for analyzing the complex interplay of forces that shape genomic architecture (Lenski et al. 1999, 2003; Wilke et al. 2001; Yedid and Bell 2001, 2002; Wilke and Adami 2002; Ofria et al. 2003, Chow et al. 2004; Edlund and Adami 2004; Ofria and Wilke 2004; Misevic et al. 2004, 2005; Ostrowski et al. 2006). Tools are available that allow such thorough analyses as testing the fitness effects of all possible single mutations in many genetic backgrounds, as we did in this study to examine genomic robustness. Large-scale experiments can be performed, again as in this study that included 10 replicate populations evolved under each of 16 different combinations of population size, mutation rate, and genome-length variation. Of course, one could argue that digital organisms live in a peculiar world, one where metabolism is based on numerical computations rather than chemical reactions, and therefore conclusions cannot be extrapolated to the organic world. But as Wilke and Adami (2002) argued, rather than being a problem, this profound difference from organic life is one of the major strengths of research with digital organisms: if a principle applies to digital organisms, can be derived from mathematical models, and is seen to operate in biochemical organisms, then it represents a general principle and not an artifact of any particular system. Or as John Maynard Smith (1992) said, "So far, we have been able to study only one evolving system and we cannot wait for interstellar flight

to provide us with a second. If we want to discover generalizations about evolving systems, we will have to look at artificial ones."

ACKNOWLEDGMENTS

We thank C. Adami for valuable suggestions and B. Baer for excellent technical assistance. This research was supported by grants BMC2003-00066 and *de Madariaga* from the Spanish Ministerio de Educatión y Ciencia-Fondo Europeo de Desarrollo Regional and by the European Molecular Biology Organization Young Investigator Program to SFE, by National Science Foundation grant CCF-0523449 to CO and REL, and by the Defense Advanced Research Projects Agency "FunBio" grant (344-4065) to Simon Levin (Princeton University).

LITERATURE CITED

Adami, C. 1998. Introduction to artificial life. Springer, New York.

———. 2006. Digital genetics: unraveling the genetic basis of evolution. Nat. Rev. Genet. 7:109–118.

Barkai, N., and S. Leibler. 1997. Robustness in simple biochemical networks. Nature 387:913–917.

Bürger, R. 2000. The mathematical theory of selection, recombination, and mutation. Wiley, Chichester, U.K.

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. A. G. M., J. Hermisson, G. P. Wagner, L. Ancel Meyers, H. Bagheri-Chaichian, J. L. Blanchard, L. Chao, J. M. Cheverud, S. F. Elena, W. Fontana, G. Gibson, T. F. Hansen, D. Krakauer, R. C. Lewontin, C. Ofria, S. H. Rice, G. von Dassow, A. Wagner, and M. C. Whitlock. 2003. Evolution and detection of genetic robustness. Evolution 57:1959–1972.

Earl, D. J., and M. W. Deem. 2004. Evolvability is a selectable trait. Proc. Natl. Acad. Sci. USA 101:11531–11536.

Edlund, J. A., and C. Adami. 2004. Evolution of robustness in digital organisms. Artif. Life 10:167–179.

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

———. 2001. Epistasis between new mutations and genetic background and a test of genetic canalization. Evolution 55:1746–1752.

Gibson, G., and G. P. Wagner. 2000. Canalization in evolutionary genetics: a stabilizing theory? BioEssays 22:372–380.

Hermisson, J., and G. P. Wagner. 2004. The population genetic theory of hidden variation and genetic robustness. Genetics 168:2271–2284.

Huynen, M. A., P. F. Stadler, and W. B. Fontana. 1996. Smoothness within ruggedness: the role of neutrality in adaptation. Proc. Natl. Acad. Sci. USA 93:397–401.

Kawecki, T. J. 2000. The evolution of genetic canalization under fluctuating selection. Evolution 54:1–12.

Kimura, M. 1962. On the probability of fixation of mutant genes in a population. Genetics 47:713–719.

Kondrashov, A. S. 1982. Selection against harmful mutations in large sexual and asexual populations. Genet. Res. 40:325–332.

Krakauer, D. C., and J. B. Plotkin. 2002. Redundancy, antiredundancy and the robustness of genomes. Proc. Natl. Acad. Sci. USA 99:1405–1409.

Lenski, R. E. 2001. Twice as natural. Nature 414:255.

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.

Lenski, R. E., J. E. Barrick, and C. Ofria. 2006. Balancing robustness and evolvability. PLoS Biol. 4:(12)e428.

Maynard Smith, J. 1992. Byte-sized evolution. Nature 335:772–773.

Misevic, D., R. E. Lenski, and C. Ofria. 2004. Sexual reproduction and Muller's ratchet in digital organisms. Pp. 340–345 *in* J. Pollack, M. Bedau, P. Husbands, T. Ikegami, and R. A. Watson, eds. Artificial life IX. MIT Press, Cambridge, MA.

Misevic, D., C. Ofria, and R. E. Lenski. 2005. Sexual reproduction reshapes the genetic architecture of digital organisms. Proc. R. Soc. Lond. B. 273:457–464.

Nowak, M. A., M. C. Boerlijst, J. Cooke, and J. Maynard Smith. 1997. Evolution of genetic redundancy. Nature 388:167–171.

Ofria, C., C. Adami, and T. C. Collier. 2002. Design of evolvable computer languages. IEE Trans. Evol. Comp. 6:420–424.

———. 2003. Selective pressures on genomes in molecular evolution. J. Theor. Biol. 222:477–483.

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

O'Neill, B. 2003. Digital evolution. PLoS Biol. 1:11–14.

Ostrowski, E. A., C. Ofria, and R. E. Lenski. 2007. Ecological specialization and adaptive decay in digital organisms. Am. Nat. 169:E1–E20.

Proulx, S. R., and P. C. Phillips 2005. The opportunity for canalization and the evolution of genetic networks. Am. Nat. 165:147–162

Ray, T. S. 1991. An approach to the synthesis of life. Pp. 371–408 *in* C. G. Langton, C. Taylor, J. D. Farmer, and S. Rasmussen, eds. Artificial life II. Addison-Wesley, Redwood City, CA.

Scharloo, W. 1991. Canalization: genetic and developmental aspects. Annu. Rev. Ecol. Syst. 22:65–93.

Schuster, P., and J. Swetina. 1988. Stationary mutant distributions and evolutionary optimization. Bull. Math. Biol. 50:635–660.

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

Stearns, S. C., and T. J. Kawecki. 1994. Fitness sensitivity and the canalization of life-history traits. Evolution 48:1438–1450.

van Nimwegen, E., J. P. Crutchfield, and M. Huynen. 1999. Neutral evolution of mutational robustness. Proc. Natl. Acad. Sci. USA 96:9716–9720.

Wagner, A. 1999. Redundant gene functions and natural selection. J. Evol. Biol. 12:1–16.

———. 2005. Robustness and evolvability in living systems. Princeton Univ. Press, Princeton, NJ.

Wagner, A., and P. F. Stadler. 1999. Viral RNA and evolved mutational robustness. J. Exp. Zool. 285:119–127.

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

Wagner, G. P., G. Booth, and H. Bagheri-Chaichian. 1997. A population genetic theory of canalization. Evolution 51:329–347.

Wilke, C. O. 2001a. Adaptive evolution on neutral networks. Bull. Math. Biol. 63:715–730.

———. 2001b. Selection for fitness versus selection for robustness in RNA secondary structure folding. Evolution 55:2412–2420.

Wilke, C. O., and C. Adami. 2001. Interaction between directional epistasis and average mutational effects. Proc. R. Soc. Lond. B 268:1469–1474.

———. 2002. The biology of digital organisms. Trends Ecol. Evol. 17:528–532.

———. 2003. Evolution of mutational robustness. Mut. Res. 522:3–11.

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.

Yedid, G., and G. Bell. 2001. Evolution in an electronic microcosm. Am. Nat. 157:465–487.

———. 2002. Macroevolution simulated with autonomously replicating computer programs. Nature 420:810–812.

Associate Editor: T. Hansen