Chapter 14
Genome Size and the Extinction of Small Populations

Thomas LaBar and Christoph Adami

Abstract Although extinction is ubiquitous throughout the history of life, the factors that drive extinction events are often difficult to decipher. Most studies of extinction focus on inferring causal factors from past extinction events, but these studies are constrained by our inability to observe extinction events as they occur. Here, we use digital evolution to avoid these constraints and study "extinction in action". We examine the genetic mechanisms driving the relationship between genome size and population extinction. We find that genome expansions enhance extinction risk through two genetic mechanisms that increase a population's lethal mutational burden: an increased lethal mutation rate and an increased likelihood of stochastic reproduction errors. This result, contrary to the expectation that genome expansions should buffer mutational effects, suggests a role for epistasis in driving extinction. We discuss biological analogues of these digital "genetic" mechanisms and how large genome size may inform which natural populations are at an increased risk of extinction.

14.1 Introduction

The ubiquity of extinction events throughout the history of life [20] and the increasing realization that Earth's biosphere may be experiencing a sixth mass extinction [4] drive interest in determining the factors that cause certain species, but not others, to go extinct [33]. It is accepted that genetic [38, 47], demographic [32, 35], environmental [28, 50], and ecological [9, 12, 39] factors contribute to species ex-

Thomas LaBar
Department of Microbiology and Molecular Genetics, Michigan State University e-mail: labartho@msu.edu

Christoph Adami
Department of Microbiology and Molecular Genetics, Michigan State University e-mail: adami@msu.edu

tinctions. Beyond those deterministic factors, chance events also likely influence some extinction events [41, 49]. Here, we focus on the genetic factors influencing extinction, specifically the role of small population size and genetic drift [31].

In small populations, weak purifying selection leads to increased fixation of small-effect deleterious mutations [56]. As multiple deleterious mutations fix, the absolute fitness of the population may decrease, resulting in a decrease in population size. This decreased population size further weakens selection, leading to the fixation of additional deleterious mutations and a further decrease in population size. This process continues until population extinction occurs. This positive feedback loop between decreased population size and deleterious mutation fixation is known as a mutational meltdown [29]. Mathematical models of mutational meltdowns suggest that even intermediate-sized asexual populations (approximately 10^3 to 10^4 individuals) can quickly go extinct [15, 30]. Likewise, small sexual populations are also vulnerable to fast meltdowns [25].

The concept of a mutational meltdown provides a population-genetic mechanism for extinction. However, it is still uncertain which factors beyond population size influence the likelihood of a meltdown. If deleterious mutation accumulation drives mutational meltdowns, then species with a greater genomic mutation rate should be at a greater risk of extinction [46, 61]. Genome expansions (i.e., mutations that increase genome size) are another proposed genetic mechanism that could lead to population extinction. Indeed, there is some evidence that genome size positively correlates with extinction risk in certain clades of multicellular organisms [53, 54].

While the relationship between high mutation rates and extinction suggests that larger genome size heightens extinction risk solely by increasing mutation rates, the connection between genome size and extinction can be complicated. If genome expansions lead to increased neutrality, the overall genomic mutation rate may increase, but the deleterious mutation rate will remain constant. Species with larger genomes should only face an increased mutational burden if genome expansions lead to increased genome content under purifying selection. For example, potential detrimental molecular interactions between an original genomic region and its duplicate may result in an increased mutational burden [11]. As genome expansions are likely to lead to many alterations in the distribution of mutational effects, it is still unclear which genetic mechanisms lead genome expansions to drive population extinction.

It is difficult to test the role of genome size in extinction in both natural and laboratory model systems. Here, we use digital experimental evolution [2, 5, 18, 22, 40] to test whether genome expansions can drive population extinction. In a previous study with the digital evolution system Avida [37] on the role of population size in the evolution of complexity, we found that the smallest populations evolved the largest genomes and the most novel traits, but also had the greatest extinction rates [23]. Now, we use Avida to test explicitly the mechanisms behind the role of genome size in the extinction of small populations.

Avida differs from previous models of extinction in small populations in the mode of selection. Unlike mutational meltdown models [31], where selection is hard and the accumulation of deleterious mutations directly leads to population ex-

tinction, selection is primarily soft in Avida and deleterious mutations alter relative fitness (i.e., competitive differences between genotypes), not absolute fitness (i.e., differences in the number of viable offspring between genotypes). Extinction occurs in Avida through "lethal," or "non-viable," mutations that prevent their bearer from reproducing. These non-viable avidians occupy a portion of the limited space allocated to an avidian population, thus reducing the effective population size and potentially causing extinction over time.

We find multiple genetic mechanisms lead genome expansions to drive the extinction of small populations. Increased genome size not only leads to an increase in the genomic mutation rate, but specifically to an increase in the lethal mutation rate. Elevated lethal mutation rates in large-genome genotypes are likely due to detrimental interactions between ancestral genome regions and duplicated genome content. Additionally, we show that genotypes with large genomes have an elevated probability of stochastic replication errors during reproduction (i.e., stochastic viability), further elevating the likelihood of offspring non-viability and extinction. These results suggest that large genome size does elevate the risk of population extinction due to an increased lethal mutational burden from multiple genetic mechanisms.

14.2 Methods

14.2.1 Avida

Here we review those aspects of Avida (version 2.14; available at `https://github.com/devosoft/avida`) relevant to the current study (see [37] for a complete overview). In Avida, simple computer programs ("avidians") compete for the resources required to undergo self-replication and reproduction. Each avidian consists of a genome of computer instructions drawn from a set of twenty-six available instructions in the Avida genetic code. A viable asexual avidian genome must contain the instructions to allocate a new (offspring) avidian genome, copy the instructions from the parent genome to the offspring genome, and divide the offspring genome to produce a new avidian. During this copying process, mutations may occur, introducing heritable variation into the population. This genetic variation causes phenotypic variation: avidians with different genomes may self-replicate at different speeds. As faster self-replicators outcompete slower self-replicators, this heritable variation results in differential fitness between avidians. Therefore, an Avida population undergoes Darwinian evolution [1, 40]. Avida has previously been used to test hypotheses concerning the evolution of genome size [17, 23], the role of population size in evolution [13, 23, 24, 36], and the consequences of population extinction [48, 58, 59, 60].

The Avida world consists of a grid of N cells; each cell can be occupied by at most one avidian. Thus, N is the maximum population size for the Avida environment. While avidian populations are usually at carrying capacity, the presence of

lethal mutations can reduce their population size below this maximum size. Here, offspring can be placed into any cell, simulating a well-mixed environment (i.e., no spatial structure). If a cell is occupied by another avidian, the new offspring will overwrite the occupant. The random placement of offspring avidians adds genetic drift to Avida populations, as avidians are overwritten without regard to fitness.

Fitness for an avidian genotype is estimated as the ratio of the number of instructions a genotype executes per unit time to the number of instructions it needs to execute to reproduce. Therefore, there are two avenues for a population of avidians to increase fitness: 1) increase the number of instructions executed per unit time, or 2) decrease the number of instruction executions needed for self-replication. Avidian populations can increase the number of instructions executed by evolving the ability to input random numbers and perform Boolean calculations on these numbers (a "computational metabolism" [26]). They can also decrease the number of instruction executions necessary for reproduction by optimizing their replication machinery.

There are a variety of different implementations of mutations in Avida. Here, we used settings that differed from the default in order to improve our ability to analyze the causes of population extinction (see Table 14.1 for a list of changes to the default settings). Point mutations change one locus from one of the twenty-six Avida instructions to another random, uniformly chosen, instruction; these mutations occur upon division between parent and offspring. There is an equal probability that each instruction in the genome will receive a point mutation; thus, genome size determines the total genomic mutation rate. To model indels, we used so-called "slip" mutations. This mutational type will randomly select two loci in the genome and then, with equal probability, either duplicate or delete the instructions in the genome between those two loci. While the rate of indel mutations remains constant, the chance of large indel mutations increases as genome size grows. Finally, to ease our analysis, we required every offspring genotype to be identical to its parent's genotype before the above mutations were applied at division. This setting prevented the origin of deterministic "implicit" mutations that occur when certain genotypes undergo genome replication [2].

One aspect of Avida mutations that differs from traditional models of population extinction is the presence of non-viable mutations in addition to merely deleterious, but still viable, mutations. We call these mutations "lethal," but strictly speaking they do not kill their bearer. Instead, they prevent their bearer from successfully reproducing within the maximum allowed lifespan (i.e., they are non-viable). Here, we used the default maximum lifespan of $20 \times L$ instruction executions, where L is the genome size. In other words, this setting limits the number of times an avidian can cycle through their genome in an attempt to reproduce. Such a setting must exist in order to allow avidian genomes to be analyzed. Otherwise, non-reproducing avidians could be analyzed forever, as the only way to decide if an avidian can reproduce is to actually execute the code in its genome.

In Avida, it is possible to perform experiments where mutations with certain effects are prevented from appearing in a population [10]. To enable this dynamic, the Avida program analyzes the fitness of every novel genotype that enters the popula-

Table 14.1: **Notable Avida parameters changed from default value.**

Parameter	Default Value	Changed Value	Treatment
WORLD_X	60	N	All
WORLD_Y	60	1	All
BIRTH_METHOD	0	4	All
COPY_MUT_PROB	0.0075	0.0	All
DIV_MUT_PROB	0.0	μ	All
DIVIDE_INS_PROB	0.05	0.0	All
DIVIDE_DEL_PROB	0.05	0.0	All
DIVIDE_SLIP_PROB	0.0	0.01	Variable Genome Size
REQUIRE_EXACT_COPY	0	1	All
REVERT_FATAL	0.0	1.0	Lethal-reversion
REVERT_DETRIMENTAL	0.0	1.0	Deleterious-reversion

tion and, if the fitness is of the pre-set effect, the mutation is reverted. This system allows experimenters the ability to determine the relevance of certain mutational effects to evolution. However, mutations of certain effects can still enter the population if their fitness effects are stochastic. An avidian has stochastic fitness if its replication speed depends on characteristics of the random numbers it inputs in order to perform its Boolean calculations. Some stored numbers may alter the order in which certain instructions are executed or copied into an offspring's genome, thus altering fitness.

14.2.2 Experimental Design

To study the role of genome size in the extinction of small populations, we first evolved populations across a range of per-site mutation rates ($\mu = 0.01$ and $\mu = 0.1$) and population sizes ($N = \{5, 6, 7, 8, 10, 15, 20\}$ for $\mu = 0.01$ and $N = \{10, 12, 15, 16, 17, 20, 25\}$ for $\mu = 0.1$). For each combination of population size and mutation rate we evolved 100 populations for at most 10^5 generations. Each population was initialized at carrying capacity with N copies of the default Avida ancestor (which has 100 instructions) with all excess instructions removed; this resulted in an ancestor with a genome of 15 instructions (only those needed for replication). By using an ancestral genotype with an almost-minimal genome (avidian genotypes with smaller genomes do exist [3, 7]), we were better able to explore the consequences of genome expansions (i.e., the ancestor is close to the theoretical lower bound on genome size). Ancestral genotypes with per-site mutation rates of $\mu = 0.01$ and $\mu = 0.1$ thus have genomic mutation rates of $U = 0.15$ and $U = 1.5$ mutations/genome/generation, respectively. Genome size mutations (indels) occured at a fixed rate of 0.01 mu-

tations/genome/generation for all treatments. Additionally, for each mutation rate and population size combination, an additional 100 populations were evolved in an environment where genome size was fixed. To directly test for the role of lethal and deleterious mutations in driving extinction, we evolved 100 populations at the low mutation rate population sizes under conditions where either lethal mutations or deleterious, but non-lethal, mutations were reverted (the "lethal-reversion" and "deleterious-reversion" treatments, respectively).

14.2.3 Data Analysis

For all evolution experiments, we saved data on the most abundant genotype every ten generations. The final saved genotype was used in all analyses here. All data represent either genotypes at most ten generations before extinction (in the case of extinct populations) or genotypes from the end of the experiment (in the case of surviving populations). In order to calculate the lethal mutation rate and other relevant statistics for a genotype, we generated every single point mutation for that genotype and measured these mutants' fitness using Avida's Analyze mode. The lethal mutation rate was estimated as $U_{\text{lethal}} = \mu \times L \times p_{\text{lethal}}$, where μ is the per-site mutation rate, L is the genome size, and p_{lethal} is the probability that a random mutation will be lethal.

14.2.3.1 Analysis of the relationship between genome expansions and changes in the lethal mutation rate

To test whether genome expansions themselves were directly responsible for the increase in the lethal mutation rate or whether the lethal mutation rate increased after evolution in response to a genome expansion, we first reconstructed the line-of-descents (LODs) for each of the one hundred genotypes evolved in a population of 20 individuals with a per-site mutation rate of 0.01 mutations/site/generation (the low mutation rate). An LOD contains every intermediate genotype from the ancestral genotype to an evolved genotype and allows us to trace how genome size evolved over the course of the experiment [26]. We reduced these LODs to only contain the ancestral genotype, the genotypes that changed genome size, the genotype immediately preceding a change in genome size, and the final genotype. We measured the genome size and the lethal mutation rate for each of these remaining genotypes. Then, we measured the relationship between the change in genome size and the change in the lethal mutation rate for genome expansions, genome reductions, and the segments of evolutionary time where genome size was constant.

14.2.3.2 Analysis of stochastic viability

In order to test the possibility that some of our populations had evolved stochastic viability, we analyzed each genotype from the $N = 5$ lethal-reversion populations and each genotype from the $N = 8$, $\mu = 0.01$, original populations. These population sizes were chosen because they had the most nearly equal number of extinct and surviving populations. We performed 1000 viability trials, where a genotype was declared non-viable if it could not reproduce. A genotype was declared stochastically-viable if the number of non-viable trials was greater than 0 and less than 1000. Otherwise, it was defined as deterministically-viable.

All data analysis beyond that using Avida's Analyze Mode was performed using the Python packages NumPy version 1.12.1 [51], SciPy version 0.19.0 [21], and Pandas version 0.20.1 [34]; figures were generated using the Python package Matplotlib version 2.0.2 [19]. All Avida scripts and data analysis scripts used here are available at https://github.com/thomaslabar/LaBarAdami_GenomeSizeExtinction.

14.3 Results

14.3.1 Large genome size increases the extinction risk of small populations

To test if large genome size enhances the probability of population extinction, we evolved populations across a range of population sizes at both high (1.5 mutations/genome/generation) and low (0.15 mutations/genome/generation) mutation rates with either a fixed genome size or a variable genome size. Based on our previous work with a similar experimental setup [23], we predicted that our smallest populations would go extinct at high rates if genome size could vary (which, based on this previous study, results in genome expansion and large genome size). As expected, under the low mutation rate regime populations with variable genome sizes had greater rates of extinction than those with fixed small genomes (Fig. 14.1A). However, under the high mutation rate regime, there was no significant difference between populations with a variable genome size and populations with a constant genome size (Fig. 14.1A). Estimations of the time to extinction further support these trends: in the low mutation regime, populations where genome size could evolve went extinct in fewer generations than those where genome size was constant. There were no differences in the high mutation rate regime (Fig. 14.1B).

Next, we compared the final evolved genome size between genotypes from extinct populations and surviving populations under the Variable Genome Size treatment. Across the range of population sizes for which at least 10 populations both survived and went extinct, "extinct" genotypes evolved larger genomes than those "surviving" genotypes in the low mutation rate regime (Fig. 14.2A). In the high mutation rate regime, one population size ($N = 15$ individuals) led to surviving popu-

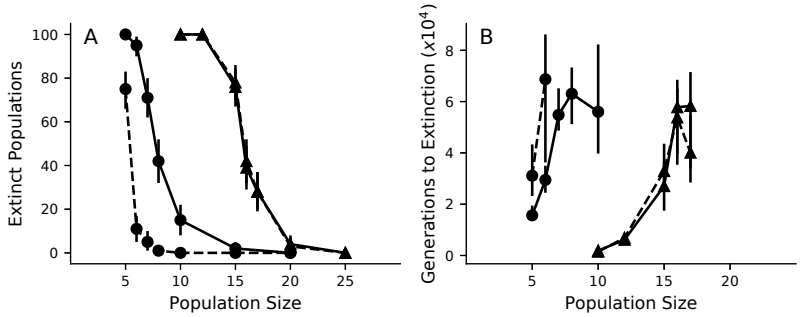

Fig. 14.1: **Possibility of genome expansions increases extinction in low mutation rate populations.** A) Number of extinct populations (out of 100) as a function of population size. Solid (dashed) lines represent variable (fixed) genome size populations. Circles (triangles) represent low (high) mutation rate populations. Error bars are bootstrapped 95% confidence intervals (10^4 samples). B) Median time to extinction for population size and mutation rate combinations. Lines and symbols same as in panel A. Error bars are bootstrapped (10^4 samples) 95% confidence intervals of the median. Data only shown for those treatments that resulted in at least ten extinct populations.

lations evolving larger genomes, while there was no statistically-significant difference for the other population sizes (Fig. 14.2B). Together, these results suggest that genome expansions and large genome size can enhance the risk of small population extinction if the initial mutation rate is too low for extinction to otherwise occur. We next focus on examining the mechanism behind the relationship between genome size and extinction in the low mutation rate populations.

14.3.2 Extinction and large genome size are associated with increases in the lethal mutational load

Avidian populations only face population-size reductions through one mechanism: parent avidians produce non-viable offspring that replace viable avidians. In other words, the lethal mutational load should drive population extinction. It is therefore possible that the increased genomic mutation rate that co-occurs with genome expansions specifically increased the genomic lethal mutation rate. The elevated lethal mutation rate then leads to an increased rate of population extinction. We first tested whether larger genomes had increased lethal mutation rates. Genome size was correlated with the lethal mutation rate across genotypes from all population sizes, supporting the hypothesis that increases in genome size result in increased lethal mutational loads and eventually population extinction (Fig. 14.3A; Spearman's $\rho \approx 0.75$, $n = 616$, $p = 1.77 \times 10^{-148}$). Next, we examined whether populations that went extinct had previously evolved greater lethal mutation rates than surviving popula-

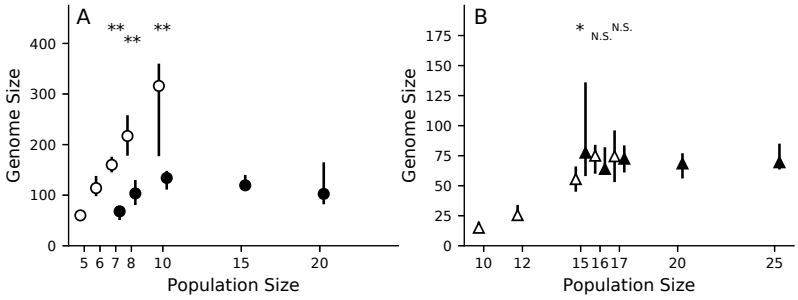

Fig. 14.2: **Extinct populations evolved larger genomes.** A) Final genome size for the low mutation rate populations from the Variable Genome Size treatment as a function of population size. Populations that survived are shown in black; populations that went extinct are shown in white. Data points are median values and error bars are bootstrapped (10^4 samples) 95% confidence intervals of the median. Data points are offset for clarity. ** indicates $p < 10^{-4}$, * indicates $p < 10^{-2}$, and N.S. indicates $p > 0.05$ for the Mann-Whitney U-test. Population sizes where fewer than ten populations went extinct (or survived) not shown. B) Final genome size for the high mutation rate populations from the Variable Genome Size treatment as a function of population size. Description same as in panel A. Population sizes where fewer than ten populations went extinct (or survived) not shown.

tions. As with the trend for genome size, extinct populations evolved greater lethal mutations rates than surviving populations (Fig. 14.3B).

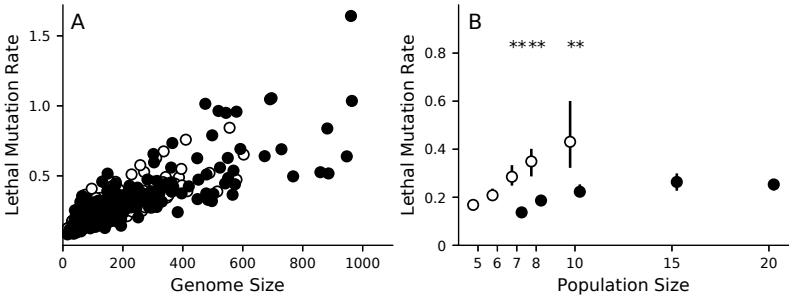

Fig. 14.3: **Lethal mutation rate correlates with genome size and population extinction.** A) The lethal mutation rate as a function of genome size for the final genotypes from each evolved low mutation rate population. B) The lethal mutation rate for extinct and surviving populations across population sizes. Error bars are bootstrapped (10^4 samples) 95% confidence intervals of the median. Population sizes where fewer than ten populations went extinct (or survived) not shown. Colors and symbols same as Fig. 14.2.

The previous data support the hypothesis that genome expansions drive population extinction by increasing the lethal mutation rate and thus the lethal mutational load. However, it is unclear whether genome expansions themselves increase the likelihood of lethal mutations (suggesting epistasis between genome expansions and the ancestral genome) or whether genome expansions merely potentiate future increases in the lethal mutation rate (due to subsequent adaptation). If genome expansions themselves increased the likelihood of lethal mutations, we expect that mutations that increase genome size should, on average, increase the lethal mutation rate. If genome expansions merely allow for the future accumulation of additional mutations that themselves increase the lethal mutation rate, there should be no relationship between mutations that increase genome size and the lethal mutation rate.

To test these two scenarios, we examined the evolutionary histories (i.e., lines-of-descent or LODs) for all $N = 20$ low mutation-rate populations. We then examined the relationship between changes in genome size and changes in the lethal mutation rate (Fig. 14.4A). When genome size was constant, the lethal mutation rate did not change on average (median change = 0.0 mutations/genome/generation). Genome size increases on average increased the lethal mutation rate (median change = 0.01 mutations/genome/generation), while genome size decreases on average decreased the lethal mutation rate (median change = -0.013 mutations/genome/generation). Additionally, the change in genome size positively correlates with the change in the lethal mutation rate (Fig. 14.4B; Spearman's $\rho = 0.67$, $n = 3600$, $p \approx 0.0$), suggesting that genome expansions directly lead to increases in the genomic lethal mutation rate.

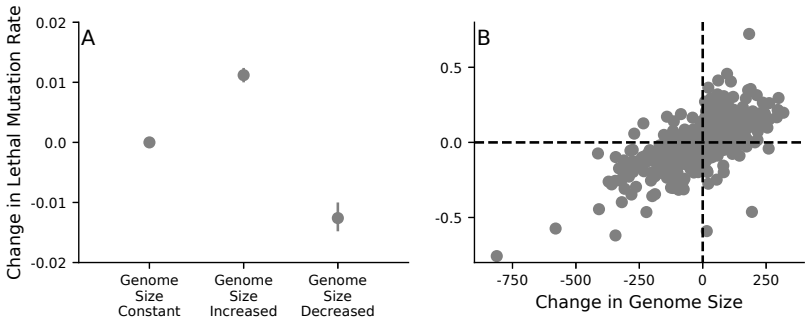

Fig. 14.4: **Insertions and deletions directly change the lethal mutation rate.** A) Change in the lethal mutation rate as a function of a mutation's effect on genome size. Each circle is the median value of all genome size alterations of a given type and error bars are bootstrapped (10^4 samples) 95% confidence intervals of the median. B) Relationship between a mutation's change in genome size and the change in the lethal mutation rate. Data same as in panel A. Dashed lines represent no change. Data points comparing genotypes with equal genome size were excluded.

14.3.3 Lethal mutation rates and stochastic viability drive population extinction

Finally, to establish the role of the lethal mutation rate in driving population extinction, we performed additional evolution experiments to test whether the absence of lethal mutations would prevent population extinction. We repeated our initial experiments (Fig. 14.1), except offspring with lethal mutations were reverted to their parental genotype (lethal-reversion treatment; see Methods for details). We also did the same experiment where deleterious, but non-lethal, mutations were reverted in order to test if deleterious mutations contributed to extinction. When populations evolved without deleterious mutations, extinction rates were similar to, if not greater than, those for populations that evolved with deleterious mutations (Fig. 14.5A). Populations that evolved with fixed-size genomes and without lethal mutations never went extinct, demonstrating how the lack of lethal mutations can prevent extinction (Fig. 14.5B). However, when these populations evolved with variable genome sizes, extinction still occurred, although at a lower rate than when lethal mutations were present (Fig. 14.5B).

While these data demonstrate that lethal mutations do primarily drive extinction risk, the fact that extinction can still occur presumably without lethal mutations is unexpected and indicates that there is a second factor that relates genome size to extinction. This is surprising, as lethal mutations are the only direct mechanism to cause extinction in Avida. One possible explanation for extinction in the lethal-reversion populations is that mutants arise in these populations that are initially viable, but later become non-viable. In other words, these populations evolve *stochastic viability*, where characteristics of the random numbers the avidians input during their life-cycle affect their ability to reproduce. These genotypes with stochastic viability would, on occasion, not be detected as lethal mutants, and thus enter the population even when lethal mutations are reverted. As they reproduce, these stochastically-viable genotypes will input other numbers and thus become, in effect, non-viable and subsequently lead to population extinction. To check if the populations that went extinct without lethal mutations did evolve stochastic viability, we tested the viability of all 100 genotypes from the lethal-reversion, variable genome size $N = 5$ populations. We also performed the same tests with the 100 genotypes from the $N = 8$ populations that evolved with lethal mutations to see if these mutants arose in our original populations.

For both sets of genotypes, we found that some genotypes were stochastically viable (Fig. 14.5C). In fact, of the 23 genotypes from populations that went extinct in the lethal-reversion treatment, 19 displayed stochastic viability. No genotypes from surviving populations were stochastically-viable. Of the 42 genotypes from populations that went extinct among our original treatment genotypes, 8 displayed stochastic viability. Two genotypes from surviving populations were stochastically-viable. Finally, we compared the genome sizes between genotypes from the lethal-reversion genotypes that were always measured as viable and those that measured as stochastic-viable. Stochastic-viable genotypes evolved larger genomes

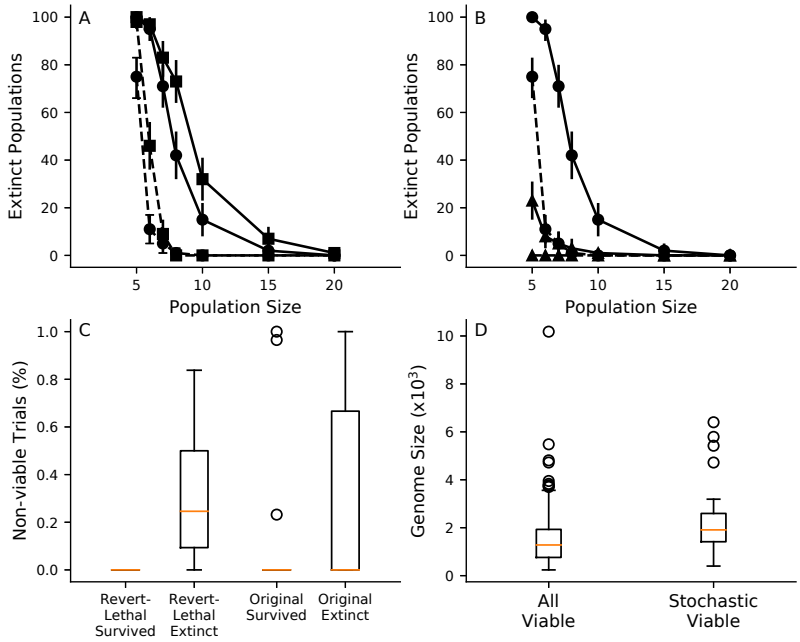

Fig. 14.5: **Evolution of stochastic viability contributes to extinction risk.** A) Number of population extinctions (out of 100 replicates) as a function of population size for the deleterious-reversion (squares) and no-reversion (circles) treatments. Dashed and solid lines represent populations from fixed genome size and variable genome size treatments, respectively. Error bars are 95% confidence intervals generated using bootstrap sampling (10^4 samples). No-reversion treatment data same as in Fig. 14.1A. B) Number of population extinctions (out of 100 replicates) as a function of population size for the lethal-reversion (triangles) and no-reversion (circles) treatments. Other symbols same as in panel A. C) Percent of viability trials (out of 1000) for which a given genotype was not viable. Values between 0 and 1 indicate stochastic viability. "Original" refers to the 100 genotypes from the $N = 8$ populations that evolved with lethal mutations. "Revert-lethal" refers the 100 genotypes from the $N = 5$ lethal-reversion populations. Red lines are median values, boxes represent the first- and third-quartile, whiskers are at-most 1.5× the relevant quartile, and circles are outliers. D) Genome size for always-viable and stochastically-viable genotypes from both the $N = 8$ no-reversion populations and the $N = 5$ lethal reversion populations.

than deterministic-viable genotypes (median = 128 instructions versus median = 191 instructions, Mann-Whitney $U = 1968.0$, $n_1 = 161$, $n_2 = 39$, $p < 2 \times 10^{-4}$; Fig. 14.5D), further suggesting that increased genome size can lead to the evolution of stochastic viability and eventual population extinction. We comment on the relevance of stochastic viability in biological populations in the Discussion below.

14.4 Discussion

We explored potential genetic mechanisms behind the relationship between genome size and the extinction of small populations. Genome expansions drive extinction because they increase the lethal mutation rate of small populations. Elevated lethal mutation rates arise through two genetic mechanisms. First, genome expansions directly increase the lethal mutation rate, suggesting that epistatic interactions between ancestral genome content and novel duplicated content lead to more lethal mutations. Second, genotypes with larger genomes have a greater likelihood of evolving stochastic viability. Both mechanisms contribute to the lethal mutational burden of small populations and together heighten the risk of population extinction.

The relationship between genome expansions and increases in the lethal mutation rate is at first counterintuitive. It is classically thought that gene/genome duplications should lead to an increase in the rate of *neutral* mutations, not lethal mutations, due to increased mutational robustness [16]. Increases in the lethal mutation rate (and not the neutral mutation rate) should only occur if there are genetic interactions (i.e., epistasis) between the ancestral genome section and the duplicated genome section. Is there evidence for gene/genome duplications leading to increased mutational load, as opposed to increased robustness? Recently, it was argued that gene duplication can also result in increased mutational fragility (not just mutational robustness) if a duplicate gene evolves to interact with its ancestral version [11]. However, more empirical studies are needed to determine whether genome expansions can elevate the mutational burden of a population to such a level that population extinction becomes a possibility.

Our second proposed mechanism underlying the connection between genome size and extinction is the evolution of stochastically-viable genotypes that can only reproduce under some environmental conditions (here, particular random number inputs). The connection behind stochastic viability and extinction in small populations is intuitive. Mutations causing stochastic viability likely have a weak effect (due to their stochastic nature) and can fix in small populations due to weakened selection. After fixation, the lethality of these mutation may be stochastically revealed and then extinction occurs. However, studies on the functional consequences of mutations responsible for extinction are rare (although see [14, 44]) and it is uncertain whether these mutations arise in populations at high extinction risk. One piece of evidence that suggests that mutations with stochastic effects might be relevant to population extinction comes from microbial experimental evolution. It has been shown that small populations have reduced extinction risk if they overexpress genes encoding molecular chaperones that assist with protein folding [45]. These overexpressed chaperones presumably compensate for other mutations that cause increased rates of stochastic protein misfolding. Therefore, mutations responsible for an increased likelihood of protein misfolding may be an example of a class of mutations with a stochastic effect that enhance extinction risk. However, this is only speculation and further work is needed to determine if stochastic viability is a possible mechanism behind extinction risk.

The most prominent model of small population extinction is the mutational meltdown model [29, 30, 31], which argues that even intermediate-sized asexual and sexual populations (i.e., 10^3 individuals) can go extinct on the order of thousands of generations. In contrast to mutational meltdown models, only very small populations go extinct in Avida, and extinction occurs on a longer timescale. The difference between our results and previous results from mutational meltdown models is likely due to differences in the character of selection between the two models. Selection is hard in mutational meltdown models, and the accumulation of deleterious mutations directly increases the probability that offspring will be non-viable [31]. In Avida, selection on deleterious mutations is soft and the accumulation of deleterious mutations is unrelated to the likelihood of non-viable offspring. Without the positive feedback loop between deleterious mutation accumulation and population size, avidian populations only evolve a high rate of non-viable mutants if they evolve large genomes, thus explaining the trends we saw here.

These differences between extinction in hard selection models and the Avida selection model emphasizes the need to consider whether selection in biological populations is primarily hard or soft. Unfortunately, there has been little resolution on this question [42, 55]. There is some evidence that soft selection may be more prevalent than hard selection. For instance, soft selection has been invoked as an explanation for why humans are able to experience high rates of deleterious mutations per generation [8, 27]. Moreover, the persistence of small, isolated populations [6, 57, 43] suggests that not only is selection primarily soft in nature, but that the extinction dynamics we study here are relevant to a subset of biological populations. While large genome size may not be the factor that causes populations to decline, it could drive an already-reduced population to extinction.

In a previous study, we observed that small populations evolved the largest genomes, the greatest phenotypic complexity, and the greatest rates of extinction [23]. This result raised the question of whether greater biological complexity itself could increase a population's rate of extinction. Although we did not test whether increased phenotypic complexity had a role in extinction, we have shown that genome size did drive small-population extinction. While it is possible that phenotypic complexity also enhanced the likelihood of extinction, the Avida phenotypic traits likely do not increase the lethal mutation rate. Thus, both high extinction rates and increased phenotypic complexity arise due to the same mechanism: greater genome size. This result illustrates an evolutionary constraint for small populations. While weakened selection and stronger genetic drift can lead to increases in biological complexity, small populations must also evolve genetic architectures that reduce the risk of extinction [45, 52]. Otherwise, small populations cannot maintain greater complexity and their lethal mutational load inexorably drives them to extinction.

Acknowledgements We thank M. Zettlemoyer and C. Weisman for comments on previous drafts of the manuscript. This work was supported in part by Michigan State University through computational resources provided by the Institute for Cyber-Enabled Research.

References

1. Adami, C.: *Introduction to Artificial Life*. TELOS, Springer Verlag (1998)
2. Adami, C.: *Digital genetics: Unravelling the genetic basis of evolution*. Nature Reviews Genetics **7**(2), 109–118 (2006)
3. Adami, C., LaBar, T.: From entropy to information: Biased typewriters and the origin of life. In: S.I. Walker, P.C. Davies, C.F. Ellis (eds.) From Matter to Life: Information and Causality, chap. 7, pp. 130–154. Cambridge University Press, Cambridge (2017)
4. Barnosky, A.D., Matzke, N., Tomiya, S., Wogan, G.O., Swartz, B., Quental, T.B., Marshall, C., McGuire, J.L., Lindsey, E.L., Maguire, K.C., et al.: *Has the Earth's sixth mass extinction already arrived?* Nature **471**(7336), 51–57 (2011)
5. Batut, B., Parsons, D.P., Fischer, S., Beslon, G., Knibbe, C.: *In silico experimental evolution: A tool to test evolutionary scenarios*. BMC Bioinformatics **14**(Suppl 15), S11 (2013)
6. Benazzo, A., Trucchi, E., Cahill, J.A., Delser, P.M., Mona, S., Fumagalli, M., Bunnefeld, L., Cornetti, L., Ghirotto, S., Girardi, M., et al.: *Survival and divergence in a small group: The extraordinary genomic history of the endangered apennine brown bear stragglers*. Proceedings of the National Academy of Sciences **114**(45), E9589–E9597 (2017)
7. CG, N., LaBar, T., Hintze, A., Adami, C.: *Origin of life in a digital microcosm*. Philosophical Transactions of the Royal Society A: Mathematical, Physical, and Engineering Sciences **375**(2109), 20160,350 (2017)
8. Charlesworth, B.: *Why we are not dead one hundred times over*. Evolution **67**(11), 3354–3361 (2013)
9. Clavero, M., García-Berthou, E.: *Invasive species are a leading cause of animal extinctions*. Trends in Ecology and Evolution **20**(3), 110–110 (2005)
10. Covert, A.W., Lenski, R.E., Wilke, C.O., Ofria, C.: *Experiments on the role of deleterious mutations as stepping stones in adaptive evolution*. Proceedings of the National Academy of Sciences **110**(34), E3171–E3178 (2013)
11. Diss, G., Gagnon-Arsenault, I., Dion-Coté, A.M., Vignaud, H., Ascencio, D.I., Berger, C.M., Landry, C.R.: *Gene duplication can impart fragility, not robustness, in the yeast protein interaction network*. Science **355**(6325), 630–634 (2017)
12. Dunn, R.R., Harris, N.C., Colwell, R.K., Koh, L.P., Sodhi, N.S.: *The sixth mass coextinction: are most endangered species parasites and mutualists?* Proceedings of the Royal Society of London B: Biological Sciences **276**(1670), 3037–3045 (2009)
13. Elena, S.F., Wilke, C.O., Ofria, C., Lenski, R.E.: *Effects of population size and mutation rate on the evolution of mutational robustness*. Evolution **61**(3), 666–674 (2007)
14. Fry, E., Kim, S.K., Chigurapti, S., Mika, K.M., Ratan, A., Dammermann, A., Mitchell, B.J., Miller, W., Lynch, V.J.: *Accumulation and functional architecture of deleterious genetic variants during the extinction of Wrangel Island mammoths*. bioRxiv p. 137455 (2017)
15. Gabriel, W., Lynch, M., Burger, R.: *Muller's ratchet and mutational meltdowns*. Evolution **47**(6), 1744–1757 (1993)
16. Gu, Z., Steinmetz, L.M., Gu, X., Scharfe, C., Davis, R.W., Li, W.H.: *Role of duplicate genes in genetic robustness against null mutations*. Nature **421**(6918), 63–66 (2003)
17. Gupta, A., LaBar, T., Miyagi, M., Adami, C.: *Evolution of genome size in asexual digital organisms*. Scientific Reports **6**, 25,786 (2016)
18. Hindré, T., Knibbe, C., Beslon, G., Schneider, D.: *New insights into bacterial adaptation through in vivo and in silico experimental evolution*. Nature Reviews Microbiology **10**(5), 352–365 (2012)
19. Hunter, J.D.: *Matplotlib: A 2D graphics environment*. Computing In Science & Engineering **9**(3), 90–95 (2007)
20. Jablonski, D.: *Background and mass extinctions: The alternation of macroevolutionary regimes*. Science **231**(4734), 129–133 (1986)
21. Jones, E., Oliphant, T., Peterson, P., et al.: SciPy: Open source scientific tools for Python (2001–)

22. Kawecki, T.J., Lenski, R.E., Ebert, D., Hollis, B., Olivieri, I., Whitlock, M.C.: *Experimental evolution*. Trends in Ecology & Evolution **27**(10), 547–560 (2012)
23. LaBar, T., Adami, C.: *Different evolutionary paths to complexity for small and large populations of digital organisms*. PLoS Computational Biology **12**(12), e1005,066 (2016)
24. LaBar, T., Adami, C.: *Evolution of drift robustness in small populations*. Nature Communications **8**, 1012 (2017)
25. Lande, R.: *Risk of population extinction from fixation of new deleterious mutations*. Evolution **48**(5), 1460–1469 (1994)
26. Lenski, R.E., Ofria, C., Pennock, R.T., Adami, C.: *The evolutionary origin of complex features*. Nature **423**(6936), 139–144 (2003)
27. Lesecque, Y., Keightley, P.D., Eyre-Walker, A.: *A resolution of the mutation load paradox in humans*. Genetics **191**(4), 1321–1330 (2012)
28. Lindsey, H.A., Gallie, J., Taylor, S., Kerr, B.: *Evolutionary rescue from extinction is contingent on a lower rate of environmental change*. Nature **494**(7438), 463–467 (2013)
29. Lynch, M., Bürger, R., Butcher, D., Gabriel, W.: *The mutational meltdown in asexual populations*. Journal of Heredity **84**(5), 339–344 (1993)
30. Lynch, M., Conery, J., Burger, R.: *Mutation accumulation and the extinction of small populations*. American Naturalist **146**(4), 489–518 (1995)
31. Lynch, M., Gabriel, W.: *Mutation load and the survival of small populations*. Evolution **44**(7), 1725–1737 (1990)
32. Matthies, D., Bräuer, I., Maibom, W., Tscharntke, T.: *Population size and the risk of local extinction: empirical evidence from rare plants*. Oikos **105**(3), 481–488 (2004)
33. Maynard Smith, J.: *The causes of extinction*. Philosophical Transactions of the Royal Society of London B: Biological Sciences **325**(1228), 241–252 (1989)
34. McKinney, W.: *Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython*. O'Reilly Media, Inc. (2012)
35. Melbourne, B.A., Hastings, A.: *Extinction risk depends strongly on factors contributing to stochasticity*. Nature **454**(7200), 100–103 (2008)
36. Misevic, D., Lenski, R.E., Ofria, C.: *Sexual reproduction and Muller's ratchet in digital organisms*. In: Ninth International Conference on Artificial Life, pp. 340–345 (2004)
37. Ofria, C., Bryson, D.M., Wilke, C.O.: Avida: A software platform for research in computational evolutionary biology. In: M. Komosinski, A. Adamatzky (eds.) Artificial Life Models in Software, pp. 3–35. Springer London (2009)
38. O'Grady, J.J., Brook, B.W., Reed, D.H., Ballou, J.D., Tonkyn, D.W., Frankham, R.: *Realistic levels of inbreeding depression strongly affect extinction risk in wild populations*. Biological Conservation **133**(1), 42–51 (2006)
39. Pedersen, A.B., Jones, K.E., Nunn, C.L., Altizer, S.: *Infectious diseases and extinction risk in wild mammals*. Conservation Biology **21**(5), 1269–1279 (2007)
40. Pennock, R.T.: *Models, simulations, instantiations, and evidence: the case of digital evolution*. Journal of Experimental & Theoretical Artificial Intelligence **19**(1), 29–42 (2007)
41. Raup, D.M.: *Extinction: Bad genes or bad luck?* WW Norton & Company, New York (1992)
42. Reznick, D.: *Hard and soft selection revisited: How evolution by natural selection works in the real world*. Journal of Heredity **107**(1), 3–14 (2015)
43. Robinson, J.A., Ortega-Del Vecchyo, D., Fan, Z., Kim, B.Y., Marsden, C.D., Lohmueller, K.E., Wayne, R.K., et al.: *Genomic flatlining in the endangered island fox*. Current Biology **26**(9), 1183–1189 (2016)
44. Rogers, R.L., Slatkin, M.: *Excess of genomic defects in a woolly mammoth on Wrangel Island*. PLoS Genetics **13**(3), e1006,601 (2017)
45. Sabater-Muñoz, B., Prats-Escriche, M., Montagud-Martínez, R., López-Cerdán, A., Toft, C., Aguilar-Rodríguez, J., Wagner, A., Fares, M.A.: *Fitness trade-offs determine the role of the molecular chaperonin GroEL in buffering mutations*. Molecular Biology and Evolution **32**(10), 2681–2693 (2015)
46. Singh, T., Hyun, M., Sniegowski, P.: *Evolution of mutation rates in hypermutable populations of Escherichia coli propagated at very small effective population size*. Biology Letters **13**(3), 20160,849 (2017)

47. Spielman, D., Brook, B.W., Frankham, R.: *Most species are not driven to extinction before genetic factors impact them*. Proceedings of the National Academy of Sciences **101**(42), 15,261–15,264 (2004)
48. Strona, G., Lafferty, K.D.: *Environmental change makes robust ecological networks fragile*. Nature Communications **7**, 12,462 (2016)
49. Turner, C.B., Blount, Z.D., Lenski, R.E.: *Replaying evolution to test the cause of extinction of one ecotype in an experimentally evolved population*. PloS One **10**(11), e0142,050 (2015)
50. Urban, M.C.: *Accelerating extinction risk from climate change*. Science **348**(6234), 571–573 (2015)
51. Van Der Walt, S., Colbert, S.C., Varoquaux, G.: *The NumPy array: A structure for efficient numerical computation*. Computing in Science & Engineering **13**(2), 22–30 (2011)
52. Vermeij, G.J., Grosberg, R.K.: *Rarity and persistence*. Ecology Letters **21**(1), 3–8 (2018)
53. Vinogradov, A.E.: *Selfish DNA is maladaptive: Evidence from the plant Red List*. Trends in Genetics **19**(11), 609–614 (2003)
54. Vinogradov, A.E.: *Genome size and extinction risk in vertebrates*. Proceedings of the Royal Society of London B: Biological Sciences **271**, 1701–1706 (2004)
55. Wallace, B.: *Fifty Years of Genetic Load: An Odyssey*. Cornell University Press (1991)
56. Whitlock, M.C., Bürger, R., Dieckmann, U.: Fixation of new mutations in small populations. In: R. Ferrire, U. Dieckmann, D. Couvet (eds.) Evolutionary Conservation Biology, chap. 9, pp. 155–170. Cambridge University Press, Cambridge (2004)
57. Xue, Y., Prado-Martinez, J., Sudmant, P.H., Narasimhan, V., Ayub, Q., Szpak, M., Frandsen, P., Chen, Y., Yngvadottir, B., Cooper, D.N., et al.: *Mountain gorilla genomes reveal the impact of long-term population decline and inbreeding*. Science **348**(6231), 242–245 (2015)
58. Yedid, G., Ofria, C., Lenski, R.: *Historical and contingent factors affect re-evolution of a complex feature lost during mass extinction in communities of digital organisms*. Journal of Evolutionary Biology **21**(5), 1335–1357 (2008)
59. Yedid, G., Ofria, C.A., Lenski, R.E.: *Selective press extinctions, but not random pulse extinctions, cause delayed ecological recovery in communities of digital organisms*. The American Naturalist **173**(4), E139–E154 (2009)
60. Yedid, G., Stredwick, J., Ofria, C.A., Agapow, P.M.: *A comparison of the effects of random and selective mass extinctions on erosion of evolutionary history in communities of digital organisms*. PloS One **7**(5), e37,233 (2012)
61. Zeyl, C., Mizesko, M., De Visser, J.: *Mutational meltdown in laboratory yeast populations*. Evolution **55**(5), 909–917 (2001)