Population bottleneck has only marginal effect on fitness evolution and its repeatability in dioecious Caenorhabditis elegans

Bisschop, K.; Blankers, T.; Mariën, J.; Wortel, M.T.; Egas, M.; Groot, A.T.; Visser, M.E.; Ellers, J.

DOI
10.1111/evo.14556

Publication date
2022

Document Version
Final published version

Published in
Evolution

doi:10.1111/evo.14556

Population bottleneck has only marginal effect on fitness evolution and its repeatability in dioecious *Caenorhabditis elegans*

Karen Bisschop,[1,2,3,4] Thomas Blankers,[1,2,5] Janine Mariën,[6] Meike T. Wortel,[7] Martijn Egas,[1] Astrid T. Groot,[1] Marcel E. Visser,[8] and Jacintha Ellers[6]

[1] *Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Amsterdam 1090 GE, The Netherlands*

[2] *Origins Center, Groningen, The Netherlands*

[3] *Terrestrial Ecology Unit, Ghent University, Ghent 9000, Belgium*

[4] *Laboratory of Aquatic Biology, KU Leuven Kulak, Kortrijk 8500, Belgium*

[5] *E-mail: thomasblankers@gmail.com*

[6] *Animal Ecology, VU Amsterdam, Amsterdam 1081 HV, The Netherlands*

[7] *Swammerdam Institute for Life Sciences, University of Amsterdam, Amsterdam 1090 GE, The Netherlands*

[8] *Department of Animal Ecology, Netherlands Institute of Ecology (NIOO-KNAW), Wageningen 6700 AB, The Netherlands*

Received October 19, 2021

Accepted June 22, 2022

The predictability of evolution is expected to depend on the relative contribution of deterministic and stochastic processes. This ratio is modulated by effective population size. Smaller effective populations harbor less genetic diversity and stochastic processes are generally expected to play a larger role, leading to less repeatable evolutionary trajectories. Empirical insight into the relationship between effective population size and repeatability is limited and focused mostly on asexual organisms. Here, we tested whether fitness evolution was less repeatable after a population bottleneck in obligately outcrossing populations of *Caenorhabditis elegans*. Replicated populations founded by 500, 50, or five individuals (no/moderate/strong bottleneck) were exposed to a novel environment with a different bacterial prey. As a proxy for fitness, population size was measured after one week of growth before and after 15 weeks of evolution. Surprisingly, we found no significant differences among treatments in their fitness evolution. Even though the strong bottleneck reduced the relative contribution of selection to fitness variation, this did not translate to a significant reduction in the repeatability of fitness evolution. Thus, although a bottleneck reduced the contribution of deterministic processes, we conclude that the predictability of evolution may not universally depend on effective population size, especially in sexual organisms.

KEY WORDS: *Caenorhabditis elegans*, effective population size, experimental evolution, fitness evolution, repeatability.

The predictability of organismal evolution lies at the heart of our understanding of evolutionary theory (Losos, 2017; Orr, 2005; Stern & Orgogozo, 2008) and is now rapidly gaining broader interest, as it provides the basis for emergent evolutionary forecasting applications (Blount et al., 2018; Lässig et al., 2017; Nosil et al., 2020; Wortel et al., 2021). A major question is whether evolution follows mostly deterministic trajectories, that is, repeatable in time and space, or whether nondeterministic processes dominate. Studies involving replicated populations in nature, such as repeated evolution of ecotypes, repeated colonization of islands,

Karen Bisschop and Thomas Blankers share joint first authorship.

or repeated host race formation, have provided support for both parallel and nonparallel evolutionary trajectories (Colosimo et al., 2005; Elmer et al., 2014; Losos & Ricklefs, 2009; Nosil et al., 2002). Similarly, replicated evolutionary experiments in the laboratory have shown that short-term (one or a few dozen of generations) and long-term evolution can be both repeatable and not repeatable (Barrick et al., 2020; Blount et al., 2018; Graves et al., 2017; Travisano et al., 1995). These variable outcomes likely result from different relative contributions from determinism and stochasticity: in the absence of chance events, evolution is highly repeatable (Lässig et al., 2017). It is therefore important to understand which properties of the natural world influence the balance between deterministic and stochastic processes and the extent to which this affects the repeatability of evolution.

In experimental evolution studies that test the repeatability of evolution, most adaptation is fueled by selection on standing genetic variation rather than mutation (Barrett & Schluter, 2008). Thus, the dominant deterministic force in experimental evolution is positive selection and the dominant stochastic force is genetic drift. A critical factor that determines the relative importance of selection versus drift is the effective population size. In the Wright-Fisher model, changes in allele frequencies are strongly determined by drift if the effective population size is small relative to the strength of selection (Crow & Kimura, 1970; Fisher, 1923; Wright, 1931). However, because several aspects of fitness evolution are affected by population size, theory is ambivalent about the expected relationships between population size and the effects of drift versus selection, and thus the repeatability of fitness evolution. For example, we would expect effective population size to be positively related to average fitness, as in larger populations the efficacy of selection is higher (Kimura, 1983) and the effects of drift are reduced (Kimura et al., 1963; Willi et al., 2006). However, in very large populations the multitude of genotypic combinations opens up additional evolutionary trajectories, which may decrease the repeatability relative to moderately large populations (Szendro et al., 2013). In small populations, evolutionary trajectories are more heterogeneous, because there is a smaller chance that the most beneficial variant will get fixed, so that fitness effects of substitutions are more variable (De Visser & Rozen, 2005; Rozen et al., 2008). More variable effects of substitutions can allow small populations to obtain higher fitness peaks than large populations, for example, if the most beneficial variants are fixed by chance (Rozen et al., 2008; Whitlock et al., 1995).

Sex (recombination) further complicates (the repeatability of) fitness evolution, because sex can increase the likelihood of evolution toward higher fitness in both large and small populations, but high recombination rates may prevent adaptation, especially in large populations (Weissman et al., 2010). Therefore, especially for sexual populations, theoretical research indicates that the conditions under which larger populations display adaptive advantage and higher evolutionary predictability over smaller populations depend on detailed genetic knowledge of the organism under study.

In evolutionary experiments, effective population sizes are generally varied by different numbers of clonally reproducing cells or by bottlenecking ancestral populations of sexually reproducing organisms (Kawecki et al., 2012). A population bottleneck randomly selects a subset of the available genotypes and thus reduces genetic diversity; effective population sizes will remain low after a bottleneck for an extended period, because genetic diversity is lost much more quickly due to drift in small populations compared to increasing diversity due to new mutations (Kimura et al., 1963; Wright, 1931). Empirical data generally show that fitness in evolved small populations is more variable (less repeatable) and on average lower compared to large populations (Lachapelle et al., 2015; Rozen et al., 2008; Weber, 1990; Wein & Dagan, 2019; Windels et al., 2021), with some exceptions (Miller et al., 2011; van Dijk et al., 2017). However, in these studies population sizes were still in the thousands of breeding/clonally reproducing individuals or effective populations sizes cannot be disentangled from census population sizes. The limited exploration of sexual species and populations with strongly reduced effective sizes leaves an important gap in our knowledge about the role of population size in the adaptive potential and repeatability of fitness evolution.

Here, we tested whether adaptation is faster and more repeatable in populations with large versus small effective sizes in an obligatory outbreeding line of the bacterivorous nematode *Caenorhabditis elegans*. We exposed replicate populations to a novel environment with a different bacterial prey and measured fitness as the population size after on week on the novel food source prior to and after 15 weeks of exposure to the novel conditions. To disentangle the impact of effective population size from the effects of census population size, we started the experiment with either 500 nematodes (at expected 1:1 sex ratio) from a large and genetically variable ancestral population or 500 nematodes derived from the same ancestral population subjected to a moderate or strong bottleneck. These experiments addressed two questions: (i) What is the effect of a population bottleneck on the average and maximum fitness after selection? (ii) What is the effect of population bottlenecks on the repeatability of fitness evolution?

Methods

STUDY SPECIES AND CREATION OF THE BOTTLENECKED POPULATIONS

We performed experimental evolution using the *C. elegans* D00 population from the Teotónio lab (IBENS, Paris), which is a

multiparent intercrossed population that is obligatorily outcrossing (Noble et al., 2017; Theologidis et al., 2014). Sex ratio in dioecious *Caenorhabditis* species and lines is expected to be 1:1 (Gray & Cutter, 2014). The D00 ancestral population was expanded on Nematode Growth Medium (NGM) (Stiernagle, 2006) plates seeded with *Escherichia coli* OP50 at 20°C and divided in aliquots. Care was taken during this phase to maintain sufficiently large population sizes. To create the bottlenecked populations, five aliquots of the starting population were thawed and expanded at 20°C with *E. coli* OP50 as a food source. After 6 days, from each expanded population five or 50 female nematodes were transferred to a separate plate for the strong bottleneck and moderate bottleneck treatments, respectively. The females were chosen randomly from all available females without visible embryos on a plate. We only selected females to avoid stochastic sampling of different numbers of males and females across replicates. These bottlenecked populations were then grown for 6 days on NGM *E. coli* at 20°C before collecting the nematodes in Eppendorf tubes. Simultaneously with this expansion, for the "no bottleneck" treatment the other five aliquots from the ancestral population were thawed and expanded on NGM *E. coli* at 20°C for 6 days (Fig. S1). After expansion, nematodes were collected and for each replicate the population density was estimated to calculate the transfer volume required to transfer 500 worms. In this way, no-bottleneck and bottleneck treatments always started with 500 nematodes in an expected 1:1 sex ratio, but for the no-bottleneck treatment, these 500 nematodes were offspring of diverse ancestral populations, whereas for the bottleneck treatments, these 500 nematodes were offspring of 50 or only five founder females. Each treatment had five replicates, each consisting of three plates (to avoid the loss of a replicate if one plate would fail due to contamination or human error) that were each initiated with 500 nematodes (Fig. S1). All populations of *C. elegans* were maintained on plates (Ø 9 cm) with ±12 mL NGM.

NOVEL CONDITIONS FOR EXPERIMENTAL EVOLUTION

During experimental evolution, *C. elegans* populations were grown on *Bacillus megaterium* (DSM No. 509). *Bacillus megaterium* is a poor food that results in impeded growth rates and, when given a choice, *C. elegans* avoids patches with *B. megaterium* (Shtonda & Avery, 2006).

In addition to the novel diet, the temperature was lowered to 16°C, which was done for experimental feasibility. The temperature reduction to 16°C may affect metabolic functions and defense pathways (Gómez-Orte et al., 2018) and therefore constituted an additional selection pressure. Moreover, 16S amplicon sequencing data from empty NGM plates revealed unexpected contamination of the plates (mainly bacteria from the genera *Serratia* and *Pseudomonas*). Even though empty plates did not reveal any visual bacterial growth at room temperature and the same plates were used for all the replicates in the different treatments, this contamination may have induced an unanticipated additional selection pressure. Because the effects of the three perturbations (novel food source, novel temperature, and plate contaminants) cannot be disentangled, they are considered together as the novel conditions.

EXPERIMENTAL SETUP

The experiment was initiated on fresh NGM plates with a lawn of *B. megaterium* and 500 nematodes per plate per replicate. Every week, each plate was replaced by a new plate while 500 nematodes were transferred by washing the plates, mixing the three plates per replicate, estimating the density of nematodes, and pipetting the necessary volume to the new plate. At the beginning (week 0) and end (week 15) of the experiment, large samples of the nematode populations were cryopreserved at −80°C until they were needed for fitness assessments.

FITNESS ASSESSMENT

As a fitness proxy for each replicate nematode population, we estimated the population size achieved after one week of growth on *B. megaterium* at 16°C (Fig. S2) starting from 500 individuals, drawn randomly from the week 0 or 15 population. Frozen week 0 and week 15 populations were thawed simultaneously and 250 μL of each population was expanded on NGM *E. coli* at 20°C for one week to create a common garden. After expansion, 500 nematodes were transferred to each of three fitness assessment plates (NGM *B. megaterium* at 16°C) for each of the replicates. The size of the population after one week of growth (i.e., the fitness proxy) was extrapolated from counts in droplets of 5 μL (Fig. S2). These extrapolations were strongly correlated to counts obtained using a flow cytometer (Fig. S3). All replicates were measured in triplicate. However, for some replicates the fitness was assessed on two or three separate fitness assessment days (leading to six or nine measurements, respectively; Table S1), which was possible as populations were frozen in several Eppendorf tubes at the same week. Additional details are available in the Supporting Information.

DATA ANALYSIS

All statistical analyses were done in R version 4.1.0 with packages "emmeans" version 1.6.1 (Lenth, 2021), "MuMIn" version 1.43.17 (Bartoń, 2020), "lawstat" version 3.4 (Gastwirth et al., 2020), and "lme4" version 1.1.27.1. (Bates et al., 2014), and plotting was done using "ggplot2" version 3.3.5 (Wickham, 2016).

EFFECT OF POPULATION BOTTLENECKS ON AVERAGE AND MAXIMUM FITNESS ACROSS REPLICATES

We tested for a significant increase in fitness (extrapolated population size after seven days of growth on *B. megaterium*) using linear mixed effect models. The dependent variable was the extrapolated population size and the fixed effects were week, treatment, and their interaction; replicate and fitness assessment day were included as random variables because the levels of these variables are random relative to the population they come from (typical for variables such as time/experimenter/measurement, etc.) and we care about their effects as a whole and not per level (Snijders & Bosker, 1999). Pairwise comparisons among fixed effect factor levels were done using Tukey's method and we performed the Levene's Test of Equality of Variances to investigate whether the variance of the fitness differed among treatments both before and after selection. We expected similar initial fitness for all treatments and lower final mean fitness and higher variance in fitness for the bottlenecked versus the no bottleneck populations.

We also investigated potential differences between treatments in the mean selection response (the difference between week 15 and week 0 per replicate). We used an Ordered Heterogeneity Test (Rice & Gainest, 1994), because we expected an order in the selection response: stronger response for the treatment without bottleneck compared to the bottlenecked populations and also a stronger response for the moderate bottleneck compared to the strong bottleneck. We also included alternative hypotheses where two treatments were equal, but different from the third treatment (Neuhäuser & Hothorn, 2006). The Ordered Heterogeneity Test was based on a Kruskal-Wallis rank sum test.

EFFECT OF POPULATION BOTTLENECKS ON THE REPEATABILITY OF FITNESS EVOLUTION

Repeatability of fitness evolution was measured and compared between the no bottleneck treatment, the moderate bottleneck treatment, and the strong bottleneck treatment following two criteria: (i) the variance among realized selection responses across replicates within treatments (lower variance implies higher repeatability) and (ii) variance partitioning among effects from selection and chance (more relative variance attributed to selection implies higher repeatability).

We compared differences in the variance of the selection response across the five replicates among treatments using an Ordered Heterogeneity Test (Neuhäuser & Hothorn, 2006; Rice & Gainest, 1994) based on Levene's Test of Equality of Variances. We expected that populations that underwent a (stronger) bottleneck had more variance in fitness across replicates.

To test whether the relative contribution from chance and selection depended on the presence and strength of a bottleneck, we partitioned variance in effects of selection (variance between before experimental evolution and after 15 weeks of experimental evolution), chance (variance among replicates), and measurement error (variance among fitness assessment measurements nested within replicate) for each of the three treatments. We fitted a nested ANOVA model and extracted the mean squares to obtain relative proportions of the mean squares for each effect. The data were unbalanced due to variation in the number of fitness assessment days done per replicate. We therefore subsampled the data within treatments for all possible combinations of fitness assessment days. This resulted in 108 datasets for the "no bottleneck" treatment ($3 \times 3 \times 3 \times 2 \times 2$ combinations of fitness assessment days) and eight datasets ($2 \times 2 \times 2 \times 1 \times 1$) for the "moderate" and "strong bottleneck" treatments. For each dataset, we calculated variance proportions, which can be compared across treatments because they come from balanced designs. We expected that populations that underwent a (stronger) bottleneck had more variance in fitness attributable to chance (drift) relative to selection.

Results

EFFECT OF POPULATION BOTTLENECKS ON AVERAGE AND MAXIMUM FITNESS ACROSS REPLICATES

Before exposure to the novel conditions, the populations under the moderate bottleneck treatment had a significantly lower fitness than the populations not exposed to a bottleneck (t-ratio $= -4.779$ and P-value < 0.0001) and the populations that underwent a strong bottleneck (t-ratio $= -7.152$ and P-value < 0.0001; Table 1, Fig. 1a). Also, the variance in fitness before selection was significantly smaller in the moderate bottleneck treatment compared to the treatment without bottleneck (Levene's test statistic $= 89.676$ and P-value < 0.0001) and the strong bottleneck treatment (Levene's test statistic $= 11.747$ and P-value $= 0.0019$).

After selection, fitness was higher in all treatments compared to the start of the experiment (Fig. 1a). However, the average fitness after 15 weeks of selection did not differ significantly among bottleneck treatments (Table 1). Similarly, no significant differences in variances in fitness across replicates were found among treatments (Ordered Heterogeneity test statistic $= 0.584$ and P-value $= 0.5599$). Both before and after evolution and across all treatments, there was variation among fitness assessment days (Fig. S4).

The average selection response, measured as the difference in fitness between the week 0 and week 15 sample of a replicate, was highest in the no bottleneck treatment (mean $= 15{,}097$ and median $= 14{,}184$), but not significantly different from the selection response in the moderate bottleneck treatment (mean $= 11{,}446$ and median $= 10{,}664$) and the strong bottleneck treatment (mean $= 11{,}541$ and median $= 5118$; Fig. 1a),

Table 1. Effect of population bottleneck after selection across replicates. We fitted a linear mixed effects model on the logarithm of the extrapolated counts. $R^2 = 0.80$; predicted $R^2 = 0.75$. (A) Model summary statistics for fixed effects. (B) Pairwise comparisons of the least square means within treatments and within weeks.

(A) Summary of Fixed Effects

	Estimate	SE	*t*-value	Approximate P ($>\|t\|$)
(Intercept, i.e., week 0, no bottleneck)	7.37	0.30	24.97	<0.0001
Week 15	2.22	0.22	10.23	<0.0001
Moderate bottleneck	–1.33	0.43	–3.10	0.0019
Strong bottleneck	0.84	0.43	1.96	0.0499
Week 15 × moderate bottleneck	0.94	0.37	2.57	0.0102
Week 15 × strong bottleneck	–0.94	0.37	–2.57	0.0105

(B) Pairwise Comparisons

Within treatments between weeks

Contrast	Estimate	SE	df	*t*-ratio	*P*-value
No bottleneck (week 0–15)	–2.22	0.24	123	–9.23	<0.0001
Moderate bottleneck (week 0–15)	–3.16	0.31	123	–10.07	<0.0001
Strong bottleneck (week 0–15)	–1.28	0.31	123	–4.07	0.0205

Within weeks between treatments

Contrast	Estimate	SE	df	*t*-ratio	*P*-value
Week 0 (no to moderate)	1.33	0.45	123	2.98	0.0347
Week 0 (no to strong)	–0.84	0.45	123	–1.88	0.1948
Week 0 (moderate to strong)	–2.18	0.36	123	–5.97	<0.0001
Week 15 (no to moderate)	0.39	0.38	123	1.04	0.5652
Week 15 (no to strong)	0.10	0.38	123	0.26	0.9635
Week 15 (moderate to strong)	–0.29	0.34	123	–0.87	0.6687

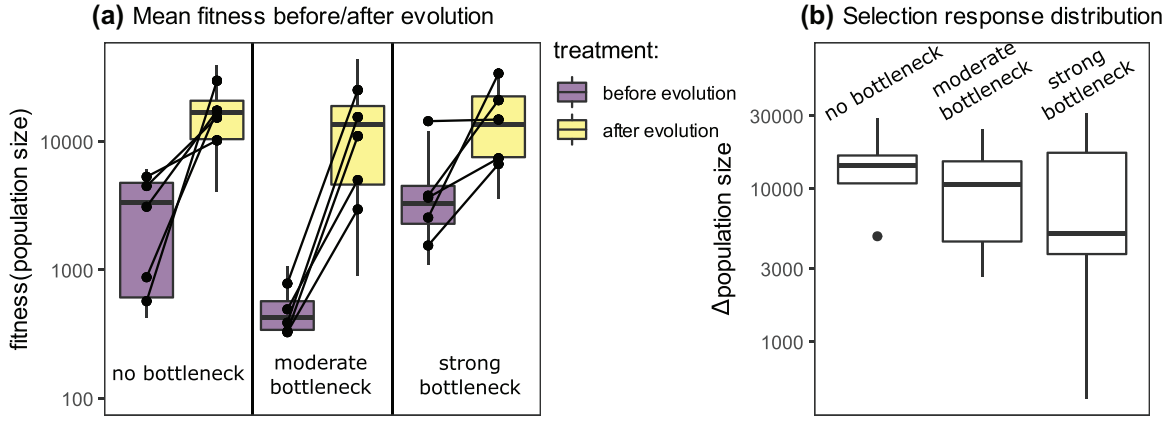

Figure 1. (a) Fitness (as population sizes measured in the fitness assay after seven days of growth on *Bacillus megaterium*) before and after evolution. Box-and-whisker plots show distributions across measurements (three measurements per assessment day, per replicate, between one and three assessment days per replicate). Black lines connect replicate averages (across measurements) before and after selection. (b) Selection response, that is, the difference in population size after seven days of growth on *B. megaterium* between week 0 (unadapted) populations and week 15 (putatively adapted) populations.

based on the Ordered Heterogeneity Test (r_sP_c statistic = 0.327, $P = 0.228$). The alternative hypothesis, that the bottleneck treatments did not differ from each other but had a lower selection response than the treatment without bottleneck, was also not significant (r_sP_c statistic = 0.573, $P = 0.073$). Neither maximum fitness or maximum selection response across replicates were lower in the moderate bottleneck or strong bottleneck treatment compared to the no bottleneck treatment (Fig. 1).

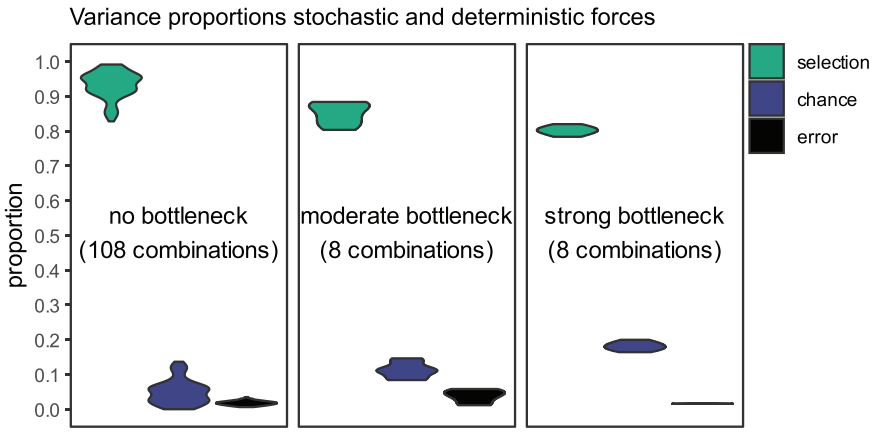

Figure 2. **Variance partitioning between selection, chance, and error. The data were subset on the level of technical replicates (measurement occasions) to balance data prior to calculating mean squares and to explore variance proportions across all possible subsets. A higher number of measurement days for the replicates of the no bottleneck scenario result in more combinations of different measurement days compared to the moderate and strong bottleneck replicates. Violin plots indicate the spread of the proportions across the different combinations.**

EFFECT OF POPULATION BOTTLENECKS ON THE REPEATABILITY OF FITNESS EVOLUTION

The variance in the response to selection was not higher in the no bottleneck treatment compared to the moderate bottleneck treatment or the strong bottleneck treatment, based on the Ordered Heterogeneity Test (r_sP_c statistic = 0.383, $P = 0.182$; Fig. 1b). The alternative hypothesis, with only a larger variance in the strong bottleneck compared to the other two treatments, was not supported either (r_sP_c statistic = 0.671, $P = 0.056$).

When we compared the distributions of relative proportions of variance explained by error, chance, and selection between the no, moderate, and strong bottleneck treatments, we found that in the strong bottleneck treatment the relative proportion of the chance effect was always larger compared to the no bottleneck treatment and the moderate bottleneck treatment (Fig. 2). Because the variance attributable to measurement error was not different among treatments, the relative variance attributable to selection was thus lower after a strong bottleneck.

Discussion

It is generally expected that large populations are better able to adapt to an environmental challenge and reach higher fitness along more predictable evolutionary trajectories compared to small populations, because selection is more efficient and drift effects are dampened in large populations (Crow & Kimura, 1970). However, predictions from theory are ambiguous, especially for sexually reproducing species, and empirical data on the repeatability of evolutionary trajectories are focused mostly on asexual unicellular species. Using an obligately outcrossing line of the nematode *C. elegans*, we found similar average and maximum fitness across three different bottleneck treatments. A larger proportion of fitness variance could be attributed to drift in populations that had experienced a strong bottleneck, but neither the presence nor magnitude of the bottleneck significantly increased the variance in fitness.

Irrespective of bottleneck treatment, we observed that nearly all replicate *C. elegans* populations evolved higher relative fitness during experimental evolution under novel conditions. Given the relatively short time of our evolutionary experiment (~21 generations) and the high genetic diversity of the *C. elegans* line used in this study (Noble et al., 2017), selection has most likely acted on standing genetic variation. Because a population bottleneck should reduce genetic variation, we had expected lower mean fitness and higher variance in fitness after evolution following a bottleneck compared to no bottleneck. These unexpected findings thus raise the question why fitness evolution in the nematodes was only marginally affected by population bottlenecks.

One potential explanation is that our bottleneck treatments did not reduce the effective population size and thus left genetic diversity unaffected. However, our fitness data support an effect of the bottleneck treatment preselection: variation in fitness between replicates as well as median fitness in the week 0 populations was lower after a bottleneck compared to populations that had no bottleneck treatment, although this was only statistically significant in the moderate bottleneck populations. Even though we cannot explain the observation of seemingly stronger effects on starting fitness in the moderate compared to the strong bottleneck result, it does not affect our conclusions from the experiments after 15 weeks. This is because (i) lower starting fitness would potentially lead to lower final

fitness and higher fitness variance, neither of which we observed, and (ii) the only significant differences we observed at the end of the experiment were between the populations without bottleneck and the populations with strong bottleneck. In addition, when we partitioned the variance, we found an increased contribution of chance effects in the populations after a strong bottleneck, which is expected when genetic diversity is reduced. Jointly, these results support a biologically relevant effect of the bottleneck treatment on the amount of genetic variation available to selection.

Another possibility is that the initial sampling of adult females for the bottlenecked treatments may have influenced the fitness results at the start. For example, sampling may have been biased toward less mobile or larger nematodes, because those nematodes are easier to pick and stand out. However, our sampling strategy likely did not favor individuals with higher fitness in the new environment for the following two reasons: (1) by avoiding females carrying embryos, we did not sample the fastest developing individuals; (2) there may be trade-offs between fitness in the ancestral environment and resilience toward novel conditions in a new environment (Haegeman et al., 2014; Muller-Landau, 2010), as well as costs of adaptation (Kassen, 2002), that suggest high fitness in one environment may not translate to high fitness in the other. Due to the sampling strategy, there is also a small chance that an adult female was not yet fertilized, which may increase the variance among replicates in the bottlenecked populations. However, because our main result is that the bottleneck had only marginal effects on both mean fitness and fitness variation, it is unlikely that our sampling strategy had a significant effect on the outcome of the experiment.

A more likely explanation for our finding that all populations evolved higher relative fitness under novel conditions is that the adaptive potential of populations with reduced genetic variation is context dependent. For example, lab studies with *E. coli* populations have shown that bottleneck effects on the repeatability of fitness evolution depend on the traits that are under selection during adaptation, for example, reduced repeatability across smaller *E. coli* populations under selection for antibiotic resistance (Windels et al., 2021) but not under selection for thermal tolerance (Wein & Dagan, 2019). Theoretical work has further shown that evolutionary predictability may not be uniformly influenced by effective population size, but that predictability is constrained in both very small and very large populations (Szendro et al., 2013). Lastly, sexual reproduction may modulate the effects of reduced genetic diversity on evolutionary repeatability in small populations (Weissman et al., 2010). Because we conducted our experiments with obligate outcrossing, sexual *C. elegans* populations, it is possible that the effects of population bottlenecks on fitness evolution were mitigated by recombination.

The genetic architecture may be a critical factor in predicting fitness evolution in relation to population bottlenecks. For example, in *C. elegans* epistatic interactions between genes underlying behavioral and fitness traits are widespread (Gaertner et al., 2012; Noble et al., 2017). Because epistasis is likely to reduce the effect of selfing on inbreeding depression (Abu Awad & Roze, 2020), these epistatic interactions may have evolved as a result of adaptation to a self-fertilizing life history with frequent cycles of exponential population growth followed by population crashes (Frézal & Félix, 2015). In addition, epigenetic changes may also contribute to the evolutionary responses observed here (Cavalli & Heard, 2019). Our common garden experimental design accounts for possible plastic and parentally (single-generation) heritable epigenetic changes, but the design does not account for any epigenetic changes that are stably inherited across multiple generations (Cavalli & Heard, 2019; Chey & Jose, 2022) and we can thus not exclude their role in driving fitness evolution. Both epistasis and epigenetic inheritance could buffer fitness evolution against reduced genetic diversity in small populations. We therefore suggest that our results fit a paradigm in which adaptive potential (and thus repeatability across replicated evolutionary events) does not unequivocally depend on effective population size, but that this relationship is shaped by the balance between diversity at neutral versus selected loci, by the strength of selection, and by the genetic architecture of selection responses (Bock et al., 2015; Carlson et al., 2014; Schrieber & Lachmuth, 2017).

In conclusion, we found that a strong population bottleneck in an obligate outcrossing line of *C. elegans* resulted in a higher contribution from drift and lower contribution from selection to fitness variation compared to populations that did not undergo a bottleneck. Importantly, due to our experimental setup we can exclude that the increased contribution from drift is due to (collinear) differences in census population size. The effects of bottlenecking on the evolution of fitness are marginal, as we observed only minor differences in fitness increase between treatments over the selection period, as well as in the repeatability of this fitness increase. Our results suggest a context-dependent relationship between genetic diversity, the effect of selection, and the predictability of evolutionary change.

ACKNOWLEDGMENTS

We would like to thank J. Teapal and the Utrecht University Large-Particle Flow Cytometry Facility (UU-LPC) for their help with the BioSorter, S. Wiezer from Aquatic Ecology at NIOO-KNAW for using the Petri plate filling machine, and S. van der Steen. We also acknowledge A. de Visser, The Predicting Evolution consortium (D. Bonte, M. Bosse, S. Declerck, M. de Vos, R. S. Etienne, S. Goossens, M. Groenen, P. Hogeweg, J. Kammenga, K. Kraaijeveld, M. Maan, F. Mortier, I. R. Pen, J. Riksen, I. Smallegange, M. van der Zee, S. van Doorn, K. Verhoeven, B. Wertheim, S. Wiezer, and L. E. Zandbergen), and the Origins

Center for helpful discussions. This work was funded by The Dutch Research Council National Science Agenda (NWA-ORC 400.17.606/4175) and a Flemish Research Foundation fellowship awarded to KB (FWO-12T5622N).

AUTHOR CONTRIBUTIONS

All authors conceived and designed the study. JM performed the experiments. KB and TB analyzed the data. KB and TB drafted the initial version of the manuscript and all authors contributed to later versions of the manuscript.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA ARCHIVING

The data and R code are available on Figshare (Bisschop et al., 2022) https://doi.org/10.21942/uva.20131868.

REFERENCES

Abu Awad, D. & Roze, D. (2020) Epistasis, inbreeding depression, and the evolution of self-fertilization. *Evolution; Internation Journal of Organic Evolution*, 74, 1301–1320.

Barrett, R.D.H. & Schluter, D. (2008) Adaptation from standing genetic variation. *Trends in Ecology & Evolution*, 23, 38–44.

Barrick, J.E., Deatherage, D.E. & Lenski, R.E. (2020) A test of the repeatability of measurements of relative fitness in the long-term evolution experiment with *Escherichia coli*. Pp. 77–89 *in* W. Banzhaf, B. H. C. Cheng, K. Deb, K. E. Holekamp, R. E. Lenski, C. Ofria, R. T. Pennock, W. F. Punch, and D. J. Whittaker, eds. Evolution in action: past, present and future: a festschrift in honor of Erik D. Goodman. Springer International Publishing, Cham, Switzerland.

Bartoń, K.A. (2020) MuMIn: multi-model inference. R package version 1.43.17.

Bates, D., Mächler, M., Bolker, B. & Walker, S. (2014) Fitting linear mixed-effects models using lme4. *J. Stat. Softw.*, 67, 1–48.

Bisschop, K., Blankers, T., Mariën, J., Wortel, T., Meike, Egas, M. & Groot, A.T., et al. (2022) Evo22-0199_DataArchive.zip. University of Amsterdam /Amsterdam University of Applied Sciences. Dataset. https://doi.org/10.21942/uva.20131868.v2

Blount, Z.D., Lenski, R.E. & Losos, J.B. (2018) Contingency and determinism in evolution: replaying life's tape. *Science*, 655.

Bock, D.G., Caseys, C., Cousens, R.D., Hahn, M.A., Heredia, S.M., Hübner, S., Turner, K.G., Whitney, K.D. & Rieseberg, L.H. (2015) What we still don't know about invasion genetics. *Molecular Ecology*, 24, 2277–2297.

Carlson, S.M., Cunningham, C.J. & Westley, P.A.H. (2014) Evolutionary rescue in a changing world. *Trends in Ecology & Evolution*, 29, 521–530.

Cavalli, G. & Heard, E. (2019) Advances in epigenetics link genetics to the environment and disease. *Nature*, 571, 489–499.

Chey, M. & Jose, A.M. (2022) Heritable epigenetic changes at single genes: challenges and opportunities in *Caenorhabditis elegans*. *Trends in Genetics*, 38, 116–119.

Colosimo, P.F., Hosemann, K.E., Balabhadra, S., Villarreal, G., Dickson, M., Grimwood, J., Schmutz, J., Myers, R.M., Schluter, D. & Kingsley, D.M. (2005) Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. *Science*, 307, 1928–1933.

Crow, J.F. & Kimura, M. (1970) An introduction to population genetics theory. New York, NY: Harper & Row.

De Visser, J.A.G.M. & Rozen, D.E. (2005) Limits to adaptation in asexual populations. *Journal of Evolutionary Biology*, 18, 779–788.

Elmer, K.R., Fan, S., Kusche, H., Spreitzer, L., M, K., A, F., Franchini, P. & Meyer, A. (2014) Parallel evolution of Nicaraguan crater lake cichlid fishes via non-parallel routes. *Nature Communication*, 5, 5168.

Fisher, R.A. (1923) XXI.—On the Dominance Ratio. *Proc. R. Soc. Edinburgh*, 42, 321–341.

Frézal, L. & Félix, M.A. (2015) *C. elegans* outside the Petri dish. *Elife*, 4, 1–14.

Gaertner, B.E., Parmenter, M.D., Rockman, M.V., Kruglyak, L. & Phillips, P.C. (2012) More than the sum of its parts: a complex epistatic network underlies natural variation in thermal preference behavior in *Caenorhabditis elegans*. *Genetics*, 192, 1533–1542.

Gastwirth, J.L., Gel, Y.R., Hui, W.L.W., Lyubchich, V., Miao, W. & Noguchi, K. (2020) Lawstat: tools for biostatistics, public policy, and law. R package.

Gómez-Orte, E., Cornes, E., Zheleva, A., Sáenz-Narciso, B., de Toro, M., Iñiguez, M., López, R., San-Juan, J.F., Ezcurra, B., Sacristán, B., et al. (2018) Effect of the diet type and temperature on the C. elegans transcriptome. *Oncotarget*, 9, 9556–9571.

Graves, J.L., Hertweck, K.L., Phillips, M.A., Han, M.V., Cabral, L.G., Barter, T.T., Greer, L.F., Burke, M.K., Mueller, L.D., Rose, M.R., et al. (2017) Genomics of parallel experimental evolution in *Drosophila*. *Molecular biology and evolution*, 34, 831–842.

Gray, J.C. & Cutter, A.D. (2014) Mainstreaming Caenorhabditis elegans in experimental evolution. *Proc. R. Soc. B Biol. Sci.*, 281, 20133055. http://doi.org/10.1098/rspb.2013.3055

Haegeman, B., Sari, T. & Etienne, R.S. (2014) Predicting coexistence of plants subject to a tolerance-competition trade-off. *Journal of Mathematical Biology*, 68, 1815–1847.

Kassen, R. (2002) The experimental evolution of specialists, generalists, and the maintenance of diversity. *Journal of Evolutionary Biology*, 15, 173–190.

Kawecki, T.J., Lenski, R.E., Ebert, D., Hollis, B., Olivieri, I. & Whitlock, M.C. (2012) Experimental evolution. *Trends in Ecology & Evolution*, 27, 547–560.

Kimura, M. (1983) The neutral theory of molecular evolution. Cambridge University Press.

Kimura, M., Maruyama, T. & Crow, J.F. (1963) The mutation load in small populations. *Genetics*, 48, 1303–1312.

Lachapelle, J., Reid, J. & Colegrave, N. (2015) Repeatability of adaptation in experimental populations of different sizes. *Proc. R. Soc. B Biol. Sci.*, 282, 20143033.

Lässig, M., Mustonen, V. & Walczak, A.M. (2017) Predicting evolution. *Nature Ecology & Evolution*, 1, 1–9.

Lenth, R. (2021) Emmeans: estimated marginal means, aka least-squares means. R package version 1.6.1.

Losos, J.B. (2017) Improbable destinies. fate, chance, and the future of evolution. New York: Riverhead Books.

Losos, J.B. & Ricklefs, R.E. (2009) Adaptation and diversification on islands. *Nature*, 457, 830–836.

Miller, C.R., Joyce, P. & Wichman, H.A. (2011) Mutational effects and population dynamics during viral adaptation challenge current models. *Genetics*, 187, 185–202.

Muller-Landau, H.C. (2010) The tolerance - fecundity trade-off and the maintenance of diversity in seed size. *Proceedings of the National Academy of Sciences*, 107, 4242–4247.

Neuhäuser, M. & Hothorn, L.A. (2006) A robust modification of the ordered-heterogeneity test. *J. Appl. Stat.*, 33, 721–727.

Noble, L.M., Chelo, I., Guzella, T., Afonso, B., Riccardi, D.D., Ammerman, P., Dayarian, A., Carvalho, S., Crist, A., Pino-Querido, A., et al.

(2017) Polygenicity and epistasis underlie fitness-proximal traits in the *Caenorhabditis elegans* multiparental experimental evolution (CeMEE) panel. *Genetics*, 207, 1663–1685.

Nosil, P., Crespi, B.J. & Sandoval, C.P. (2002) Host-plant adaptation drives the parallel evolution of reproductive isolation. *Nature*, 417, 440–443.

Nosil, P., Flaxman, S.M., Feder, J.L. & Gompert, Z. (2020) Increasing our ability to predict contemporary evolution. *Nature Communication*, 11, 1–6.

Orr, H.A. (2005) The probability of parallel evolution. *Evolution; Internation Journal of Organic Evolution*, 59, 216–220.

Rice, W.R. & Gainest, S.D. (1994) Extending nondirectional heterogeneity tests to evaluate simply ordered alternative hypotheses. *Proc. Natl. Acad. Sci. USA*, 91, 225–226.

Rozen, D.E., Habets, M.G.J.L., Handel, A. & De Visser, A.J.G.M. (2008) Heterogeneous adaptive trajectories of small populations on complex fitness landscapes. *Plos One*, 3, e1715.

Schrieber, K. & Lachmuth, S. (2017) The Genetic Paradox of Invasions revisited: the potential role of inbreeding × environment interactions in invasion success. *Biol. Rev.*, 92, 939–952.

Shtonda, B.B. & Avery, L. (2006) Dietary choice behavior in *Caenorhabditis elegans*. *Journal of Experimental Biology*, 209, 89–102.

Snijders, T.A.B. & Bosker, R. (1999) Multilevel analysis: an introduction to basic and advanced multilevel modeling. London: Sage Publications Ltd.

Stern, D.L. & Orgogozo, V. (2008) The loci of evolution: how predictable is genetic evolution?. *Evolution; Internation Journal of Organic Evolution*, 62, 2155–2177.

Stiernagle, T. (2006) Maintenance of *C. elegans*. *WormBook*,. https://doi.org/10.1895/wormbook.1.101.1.

Szendro, I.G., Franke, J., de Visser, J.A.G.M. & Krug, J. (2013) Predictability of evolution depends nonmonotonically on population size. *Proc. Natl. Acad. Sci*, 110, 571–576.

Theologidis, I., Chelo, I.M., Goy, C. & Teotónio, H. (2014) Reproductive assurance drives transitions to self-fertilization in experimental Caenorhabditis elegans. *Bmc Biology*, 12, 93.

Travisano, M., Mongold, J.A., Bennett, A.F. & Lenski, R.E. (1995) Experimental tests of the roles of adaptation, chance, and history in evolution. *Science*, 267, 87–90.

van Dijk, T., Hwang, S., Krug, J., de Visser, J.A.G.M. & Zwart, M.P. (2017) Mutation supply and the repeatability of selection for antibiotic resistance. *Physical Biology*, 14, 55005.

Weber, K.E. (1990) Increased selection response in larger populations. I. Selection for wing-tip height in Drosophila melanogaster at three population sizes. *Genetics*, 125, 579–584.

Wein, T. & Dagan, T. (2019) The Effect of population bottleneck size and selective regime on genetic diversity and evolvability in bacteria. *Genome Biol. Evol.*, 11, 3283–3290.

Weissman, D.B., Feldman, M.W. & Fisher, D.S. (2010) The rate of fitness-valley crossing in sexual populations. *Genetics*, 186, 1389–1410.

Whitlock, M.C., Phillips, P.C., Moore, F.B.G. & Tonsor, S.J. (1995) Multiple fitness peaks and epistasis. *Annu. Rev. Ecol. Syst.*, 26, 601–629.

Wickham, H. (2016) ggplot2: elegant graphics for data analysis. New York: Springer-Verlag.

Willi, Y., Buskirk, J.V. & Hoffmann, A.A. (2006) Limits to the adaptive potential of small populations. *Annual Review of Ecology, Evolution, and Systematics*, 37, 433–458.

Windels, E.M., Fox, R., Yerramsetty, K., Krouse, K., Wenseleers, T., Swinnen, J., Matthay, P., Verstraete, L., Wilmaerts, D., Van den Bergh, B., et al. (2021) Population Bottlenecks strongly affect the evolutionary dynamics of antibiotic persistence. *Molecular biology and evolution*, 38, 3345–3357.

Wortel, M.T., Agashe, D., Bailey, S.F., Bank, C., Bisschop, K., Blankers, T., Cairns, J., Sandro Colizzi, E., Cusseddu, D. & Pennings, P.S. (2021) The why, what and how of predicting evolution across biology: from disease to biotechnology to biodiversity. *EcoEvoRxiv*,. https://doi.org/10.32942/osf.io/4u3mg.

Wright, S. (1931) Evolution in mendelian populations. *Genetics*, 16, 97–159.

Associate Editor: S. Dey
Handling Editor: T. Chapman

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Figure S1: General overview figure. All treatments started from the same ancestral population that was expanded on ten petri dishes with NGM *E. coli* at 20°C. Bottlenecked populations were created from these expanded populations by transferring fifty nematodes (in blue; moderate bottleneck) or five nematodes (in orange; strong bottleneck). These populations ('before selection') were the start populations for the evolutionary experiment with fifteen weekly transfers onto fresh NGM *B. megaterium* plates at 16°C. Every time 500 nematodes were transferred per petri dish (with three petri dishes per replicate). The last generation is referred to as 'after selection'. There were five replicates per treatment.

Figure S2: Fitness assessment: setup and proxy. A) Each replicate (five replicates before selection and five replicates after selection per treatment) undergoes an expansion step of one week to obtain sufficient nematodes for the fitness assessment (i). Based on the counts of the number of nematodes after the expansion step (ii) the required volume is quantified to initiate the assessment (iii), which is done by counting the number of nematodes after seven days (iv) with three technical replicates. B) The used fitness proxy is the nematode population size after seven days.

Figure S3. Correlation between manual and BioSorter counts. (a) The x-axis shows the manual counts, while the y-axis gives the results from the flow cytometer or BioSorter. Each dot represents the same technical replicate (see Fig. S2) for which the population size was assessed. The linear regression through the origin is indicated with the solid line and the equation is presented in the upper left corner. (b) The manual counts are given on the x-axis, while the y-axis represents the difference between the estimated counts based on the equation in (a) and the obtained counts from the flow cytometer

Figure S4. Population size after seven days of growth on *B. megaterium* for week 15 (after evolution) replicates that are measured at multiple measurement days. Each box-plot shows the distribution of the three measurements taken at a given day by either observer 1 (purple, blue, yellow) or observer 2 (green). The different panels separate the three treatments.

Table S1: Number of fitness assessment days and total measurements per treatment replicate. The different columns indicate the treatments, weeks when the fitness was assessed during the experiment, the replicate numbers, fitness assessment days and total number of measurements.x