bioRxiv preprint doi: https://doi.org/10.1101/2021.10.07.463474; this version posted October 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

POPULATION GENETIC BOTTLENECK HAS ONLY MARGINAL EFFECT ON FITNESS EVOLUTION AND ITS REPEATABILITY IN DIOECIOUS *C. ELEGANS*

Bisschop Karen[*,1,2] (kbisschop.evo@gmail.com), Blankers Thomas[*,1,2] (thomasblankers@gmail.com), Mariën Janine[3] (janine.marien@vu.nl), Wortel Meike Tesse[4] (m.t.wortel@uva.nl), Egas Martijn[1] (c.j.m.egas@uva.nl), Groot Astrid T.[1] (a.t.groot@uva.nl), Visser Marcel E.[5] (m.visser@nioo.knaw.nl), Ellers Jacintha.[3] (j.ellers@vu.nl)

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

[2] Origins Center, The Netherlands

[3] Animal Ecology, VU Amsterdam, The Netherlands

[4] Swammerdam Institute for Life Sciences, University of Amsterdam, The Netherlands

[5] Department of Animal Ecology, Netherlands Institute of Ecology (NIOO-KNAW), The Netherlands

*Joint first authorship

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.07.463474; this version posted October 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

ABSTRACT

Evolution is a key process by which populations can adapt to novel conditions, but it is not well understood how predictable this process is. Predictability is expected to depend on the ratio of deterministic and stochastic processes that contribute to evolutionary change and this ratio is modulated by the 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 biased towards asexual unicellular organisms. Here, we used populations of obligately outcrossing *Caenorhabditis elegans* to test whether fitness increase and selection response were more heterogeneous after a moderate or strong population bottleneck compared to a scenario without bottleneck. Nematodes were exposed to a novel bacterial prey and lower temperature. Population sizes after one week of growth (as a proxy of fitness) were measured before and after 15 generations of evolution. We found that replicates across all (no/moderate/strong bottleneck) treatments evolved higher fitness and no significant difference in average or maximum fitness was found among treatments. Partitioning fitness variance among effects from selection and effects from chance showed that a strong (but not a moderate) bottleneck reduced the relative contribution of selection effects to fitness variation. However, the reduced contribution from selection did not translate to a significant reduction in the repeatability of fitness evolution. Thus, although a strong bottleneck reduced the contribution of deterministic evolutionary change, we found only marginal effects on quantitative measurements of repeatability in evolution. We conclude that the extent to which evolution is predictable may not universally depend on effective population size.

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.07.463474; this version posted October 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

IMPACT SUMMARY

Evolution is a key process by which populations can adapt to novel conditions. The predictability of evolution is elusive, but the extent to which evolution is predictable is central to our understanding of evolutionary processes and to emergent applications in medicine, agriculture, and conservation. Predictability is expected to be high when evolutionary change is mainly driven by selection, because this leads to repeatable evolutionary trajectories. However, chance effects, e.g. from genetic drift, reduce repeatability. In populations with smaller effective sizes (and thus reduced genetic diversity), chance effects are generally thought to be more prevalent than in large effective populations, suggesting that effective population size influences evolutionary repeatability. Recent theoretical insights cast doubt on the universality of this relationship between evolutionary repeatability and effective population size. Thus far, empirical work testing these theoretical insights is limited to asexual species, leaving a gap in our knowledge. Therefore, we performed an evolutionary experiment using sexually reproducing populations of the nematode *Caenorhabditis elegans*. Populations either went through a strong or moderate bottleneck or through no bottleneck at all to create populations with different degrees of genetic variation. These populations with different histories (but similar census size) were then exposed to novel conditions. Our results showed that neither average fitness in the novel environment or the repeatability of fitness evolution were significantly lower following a bottleneck. This was despite a reduced contribution from selection and increased contribution from chance to fitness variation in populations that experienced a strong bottleneck. We therefore show that in sexually reproducing species, increased contributions from chance in populations with lower effective sizes do not necessarily affect evolutionary repeatability. This is an important novel insight that is relevant both for our fundamental understanding as well as for evolutionary forecasting applications in agriculture and conservation, which are primarily targeted at sexual species.

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.07.463474; this version posted October 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

INTRODUCTION

The predictability of organismal evolution is a topic rapidly gaining interest, as it provides the basis for emergent evolutionary forecasting applications (Lässig et al., 2017; Wortel et al., 2021). At the heart of research into the predictability of evolution lies the question whether evolution follows mostly deterministic trajectories, and is therefore repeatable in time and space, or whether non-deterministic processes dominate. Studies involving replicated populations in nature, such as repeated evolution of ecotypes, repeated colonization of islands, or repeated host race formation, have provided support for both parallel and non-parallel evolutionary trajectories (Colosimo et al., 2005; Elmer et al., 2014; Losos and 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 balance affects the repeatability of evolution.

One core property that influences the relative contribution of selection and stochastic processes to evolutionary change is a population's ability to explore the evolutionary trajectories towards the highest fitness. We can envision evolutionary trajectories as adaptive walks through a fitness landscape, where different genotypes (through their phenotypic effects) have differential fitness (Wright, 1932). In the absence of sign epistasis, i.e. interactions among mutations causing fitness effects to depend on the genetic background, the fitness landscape is smooth: from any location in the fitness landscape there is a path that leads only upwards to a stable fitness optimum and populations will converge on the single best solution when given enough time. However, epistasis is widespread in nature and as a result empirically measured fitness landscapes are typically rugged, i.e. they have multiple, local peaks and valleys in addition to the fitness optimum (Poelwijk et al., 2007). In theory, a population may thus get stuck in a local maximum, because the only way to get to the global optimum is by mutational steps that result in lower fitness or move the population along neutral ridges, or by changes in the external conditions (De Vos et al., 2015; Gavrilets, 2004). More opportunity for selection to test genotypes (and genotype combinations) would reduce the relative contribution of chance effects and lead to more repeatable evolutionary trajectories.

All else being equal, populations of larger effective size (N_e) are expected to more efficiently explore the fitness landscape. This is because in larger populations there is more genetic variation, the efficacy of selection (Kimura, 1983) is higher, and the effects of drift are reduced compared to smaller populations (Kimura et al., 1963; Willi et al., 2006). Moreover, larger populations have a higher absolute number of mutations under constant mutation rates

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.07.463474; this version posted October 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

and mutations can be tested on more diverse genetic backgrounds in populations with higher genetic diversity. We therefore generally expect that average fitness and repeatability of fitness evolution increase with genetic variability of populations.

However, theory shows this expectation should be more nuanced. Wright (1932) already suggested that small populations may hold an evolutionary advantage over large populations because they are more likely to cross fitness valleys. More recent work supports this idea by showing that large populations are more likely to get trapped in local optima because they initially choose the path of steepest fitness increase (Jain et al., 2011), while higher stochasticity in smaller populations leads them to follow more heterogeneous adaptive trajectories (Rozen et al., 2008). In very small populations, beneficial combinations of epistatic mutations may be too rare to reliably access the trajectories towards higher fitness, resulting in increasing repeatability with increasing population size; in very large populations the multitude of combinations unlock additional trajectories in the fitness landscape, thereby decreasing the repeatability relative to moderately large populations (Szendro et al., 2013). Sex (recombination) can increase the likelihood of fitness valley crossing in both large and small populations, but high recombination rates may prevent adaptation, especially in large populations (Weissman et al., 2010). This could thus mean that both mean fitness and repeatability of fitness evolution are lower in large sexual populations with high recombination compared to small populations with low recombination. Therefore, especially for sexual populations, theoretical research indicates that the conditions under which populations of larger effective size display adaptive advantage and higher evolutionary predictability depend on detailed genetic knowledge of the organism under study.

Empirical research into the effect of effective population sizes on adaptive potential and repeatability is limited and biased towards asexual and unicellular species. In evolutionary experiments, effective population sizes are varied either 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 populations of smaller sizes is more variable (less repeatable) and on average lower compared to large populations (Lachapelle et al., 2015; Rozen et al., 2008; Wein and Dagan, 2019; Windels et al., 2021), with some exceptions (Miller et al., 2011; van Dijk et al., 2017). However, in all these studies the population size was still in the thousands of breeding/clonally reproducing individuals and cannot be disentangled from census population size. The limited exploration of sexual species and populations with strongly reduced effective sizes leaves an important gap in our knowledge about the role of effective population size in the adaptive potential and repeatability of fitness evolution.

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.07.463474; this version posted October 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Here we test whether adaptation is faster and more repeatable in populations of large versus small effective size in an obligatory outbreeding line of the bacterivorous nematode *C. elegans*. We exposed replicate populations to a novel bacterial food source (*Bacillus megaterium*) at a lower temperature (16 °C versus 20 °C) and measured fitness as the population size after one week on the novel food source prior to and after 15 weeks (at 16 °C corresponding to approximately 15 generations) of exposure to the novel conditions. In order 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 genetically variable ancestral population or 500 nematodes derived from two bottlenecked populations of the ancestor (bottleneck size was either 5 or 50 individuals). The experiments address 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? We hypothesized that - all else being equal - at least a strong population bottleneck would reduce genetic diversity, thereby limiting the potential for evolution towards higher fitness, increasing the relative contribution of stochastic effects, and reducing the repeatability of fitness evolution.

METHODS

Study species and creation of the bottlenecked populations (Fig. 1)

The bacterivorous soil nematode *Caenorhabditis elegans* is highly suitable for experimental evolution given its small body size (adult females are approximately 1 mm), high fecundity (about 300 eggs per adult female per week), short generation time (can be as short as 50h), and easiness to maintain in the lab (Gray and Cutter, 2014). Besides, frozen records can be created given that the first two larval stages can undergo cryopreservation (Gray and Cutter, 2014). The *C. elegans* D00 population from the Teotónio lab (IBENS, Paris) served as the initial source population for our experiment, which is a multiparent intercrossed population that is obligatorily outcrossing (Noble et al., 2017). Sex ratio in a dioecious *Caenorhabditis* species is expected to be 1:1 (Gray and Cutter, 2014).

The D00 source population was expanded on *E. coli* OP50 at 20°C and divided in aliquots. *E. coli* OP50 was maintained at 4°C. Bacterial lawns were added on the Nematode Growth Medium (NGM) (Stiernagle, 2006) plates by pipetting 50 µL of bacteria in L broth (10 g Bacto-tryptone, 5 g Bacto-yeast, 5 g NaCl, H_2O to 1 liter; pH 7.0). To create the bottlenecked populations, five aliquots of the starting population were defrosted by holding the cryo-tubes in the hand. When thawed, the liquid was directly pipetted onto NGM plates to prevent oxygen depletion. This initial step was performed at 20°C with *E. coli* OP50 as a food source to avoid selection during the expansion. After six days, each aliquot was used to transfer 5 female nematodes ("strong bottleneck") and 50 female nematodes ("moderate bottleneck") by hand to two separate plates (see Fig. 1). These populations could grow for six days on NGM *E. coli* at 20°C before collecting the nematodes in Eppendorf tubes. Simultaneously with this

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.07.463474; this version posted October 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

expansion, the other five aliquots from the ancestral population were thawed and pipetted onto NGM *E. coli* at 20°C for six days for the "no bottleneck" treatment (Fig. 1). After growing, nematodes were extracted from the plates by rinsing the plates with S buffer (Stiernagle, 2006) and the nematodes were collected in 10 mL tubes. The density of the populations was measured for the fifteen populations to define the necessary volume to initiate the evolutionary experiment with 500 nematodes per plate. Each treatment (no bottleneck, moderate bottleneck, and strong bottleneck) had five replicates consisting of three plates that were each initiated with 500 nematodes (Fig. 1). All populations of *C. elegans* were maintained on plates (⌀ 9 cm) with ±12 mL NGM. These NGM plates were poured using a Petri dish filling machine at NIOO-KNAW.

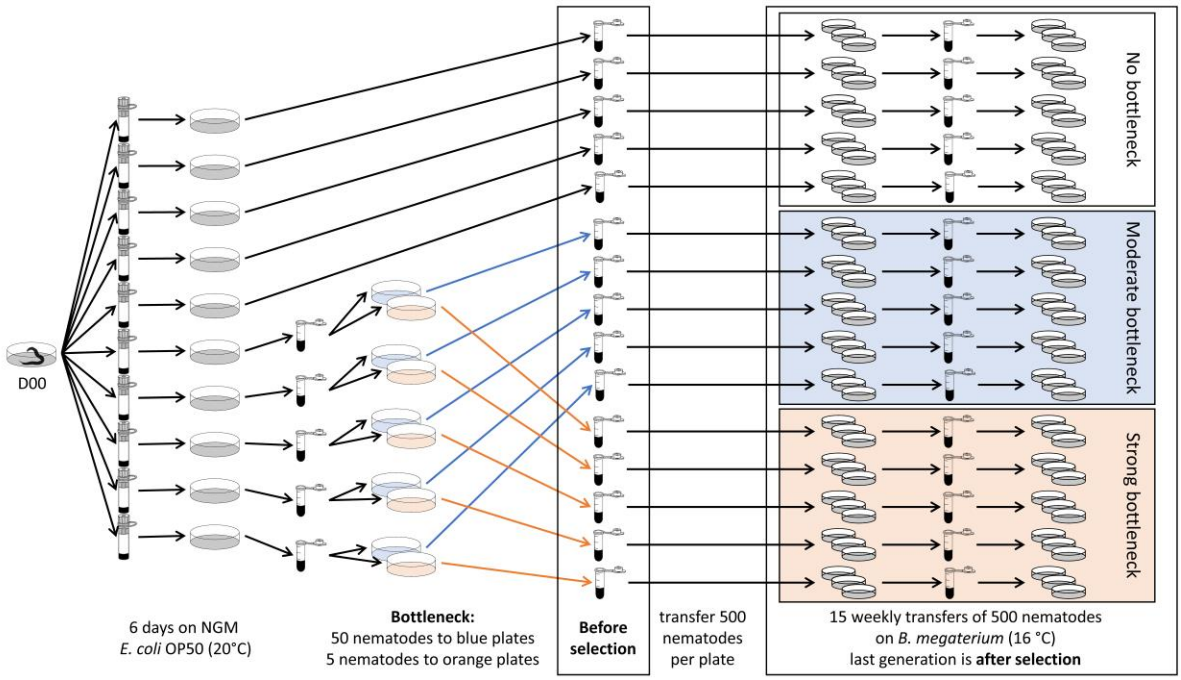


Figure 1: General overview figure. All treatments started from the same ancestral population that was expanded on ten petri dishes with NGM *E. coli* at 20°C. Bottlenecked populations were created from these expanded populations by transferring fifty nematodes (in blue; moderate bottleneck) or five nematodes (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.

Novel conditions for experimental evolution

During the experimental evolution, two novel conditions were applied: a novel bacterial resource and a lower temperature. First, the *C. elegans* populations were grown on *Bacillus megaterium* (DSM No. 509), which is a gram-positive bacterium that is known to be hard-to-eat; *C. elegans* grows slower and avoids these bacteria when given a choice (Shtonda and Avery, 2006). However, a beneficial trait of *B. megaterium* is that it contains cobalamin or

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.07.463474; this version posted October 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

vitamin B12, for which nematodes on a diet of *E. coli* have a chronic deficiency (MacNeil et al., 2013; Watson et al., 2015). *B. megaterium* was maintained at 4°C and bacterial lawns were added by pipetting 50 µL of bacteria in L broth on fresh NGM plates, which were then incubated at room temperature for 24 hours prior to transferring the worms.

Second, the growing temperature was lowered to 16°C, which was necessary to obtain about one nematode generation per week. In this way all experimental procedures could be performed on weekdays. The temperature reduction to 16°C may affect many processes such as metabolic functions and defense pathways (Gómez-Orte et al., 2018) and therefore constitute an additional selection pressure.

Third, and unexpectedly, 16S amplicon sequencing data from empty NGM plates revealed contamination of the plates (mainly bacteria from the genera *Serratia* and *Pseudomonas*). Although empty plates did not reveal any visual bacterial growth at room temperature and the same plates were used for all the treatments, this contamination may have induced an unanticipated additional selection pressure. Therefore, the effects of the three perturbations (novel food source, novel temperature, and plate contaminants) cannot be disentangled and are together considered as the novel conditions.

Experimental set-up (Fig. 1)

The experiment was initiated on fresh NGM plates with a lawn of *B. megaterium* and 500 nematodes per plate per replicate. For each replicate there were three plates that were mixed prior to the weekly transfer to sustain more genetic variation, larger population sizes, and to avoid the loss of a replicate if one plate would fail. Every week 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 plates. At the beginning (week 0) and end (week 15) of the experiment, the nematodes were cryopreserved in a 12.5% glycerol liquid freezing solution using a MrFrosty (Thermo Fisher Scientific, Waltham, MA, USA) to gradually cool them to -80°C until they were needed for fitness assessments.

Fitness assessment

As a fitness proxy for each replicate nematode population, we used the population size after one week of growth on *B. megaterium* at 16°C (Fig. 2). One week of growth is identical to a single rearing step in the evolutionary experiment as explained in the experimental set-up. An increase in the numbers of nematodes at the end of a single rearing step is thus an estimate for an increase in fitness. We assume that a higher total number of nematodes after one week of growth for the final compared to the starting population is indicative of adaptation to the novel conditions.

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.07.463474; this version posted October 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

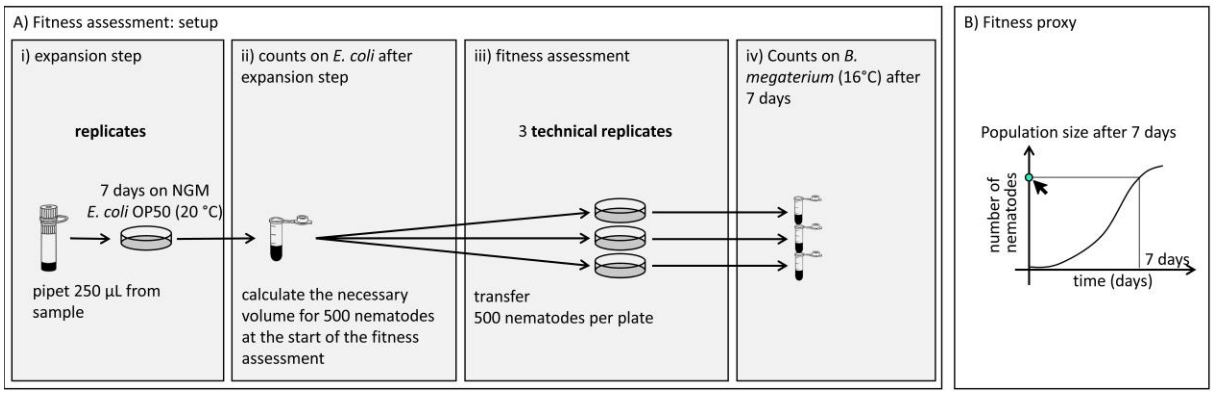


Figure 2: Fitness assessment: setup and proxy. A) Each replicate (five replicates before selection and five replicates after selection per treatment) undergoes an expansion step of one week to obtain sufficient nematodes for the fitness assessment (i). Based on the counts of the number of nematodes after the expansion step (ii) the required volume is quantified to initiate the assessment (iii), which is done by counting the number of nematodes after 7 days (iv) with three technical. B) The used fitness proxy is the nematode population size after seven days.

The set-up of the fitness assessment is presented in Fig. 2. After thawing the sample, an expansion step is performed to increase the number of nematodes. For this, 250 µL of the thawed sample is pipetted on *E. coli* at 20°C to expand the population. The expansion was for one week to create a common garden and obtain sufficient nematodes for the start of the fitness assessment. The expansion may potentially select for nematodes that (still) feed efficiently on *E. coli*, but since *E. coli* are a very accessible prey to *C. elegans* across many different wild isolates, we do not expect this to be a significant selection pressure. After expansion, the population density is estimated (by averaging the number of nematodes in three droplets of 1 µL) to calculate the necessary volume to initiate the fitness assessment with 500 nematodes. Three technical replicates are made per sample. The population size of each fitness assessment is counted by extrapolating the density estimation (number of nematodes in 5 µL). We compared our counts with the counts from a flow cytometer (BioSorter at Utrecht University) which resulted in a strong linear correlation (Fig. S1). For some populations, the fitness assessments were performed multiple times, these counts were used to quantify measurement error (Fig. S2).

Data analysis

All statistical analyses were done in R version 4.1.0 with 'glmmTMB' v. 1.1.1 (Brooks et al., 2017), 'emmeans' v. 1.6.1 (Lenth, 2021), 'MuMIn' v. 1.43.17 (Bartoń, 2020), 'lawstat' v. 3.4 (Gastwirth et al., 2020).

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 Generalized Linear Mixed Models (GLMMs). Following a goodness-of-fit of various distributions (Poisson, Negative Binomial with linear parameterization, Negative Binomial with quadratic parameterization, and Tweedie

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.07.463474; this version posted October 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

distribution), we continued with the Negative Binomial with quadratic parameterization based on the lowest AIC. The dependent variable was the extrapolated population size and the fixed effects were week (2-level categorical variable; week 0 and week 15), treatment (3-level categorical variable; no bottleneck, moderate bottleneck and strong bottleneck), and their interaction. The treatment replicates (five replicates per treatment) and fitness assessment time were considered as random variables. We performed pairwise comparisons of the least square means between treatments within weeks and between weeks within treatments and the adjustment for multiple comparisons was based on Tukey's method. Besides, 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. The variance among treatments before selection is already after the bottleneck and gives therefore a potential direct effect of bottlenecking which may influence the longer-term effect after selection.

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 and Gaines, 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 and 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 three criteria: (i) the variance among realized selection responses across replicates within treatments (lower variance implies higher repeatability), (ii) the number of replicates within each treatment that showed a significant fitness increase between samples prior to the evolutionary experiment and after 15 weeks of experimental evolution (more replicates with significant fitness increase implies higher repeatability), and (iii) variance partitioning among effects from selection and chance (more relative variance attributed to selection implies higher repeatability).

(i) We compared differences in the variance of the selection response across the five replicates among treatments using an Ordered Heterogeneity Test (Neuhäuser and Hothorn, 2006; Rice and Gaines, 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.

(ii) Significant fitness increases per replicate per treatment was assessed via GLMMs. Based on the lowest AIC score following a goodness-of-fit of various distributions (Poisson, Negative Binomial with linear parameterization, Negative Binomial with quadratic parameterization, and Tweedie distribution), we chose the Negative Binomial with quadratic parameterization. The extrapolated nematode population size was the dependent variable and

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.07.463474; this version posted October 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

the independent variables were week (2-level categorical variable; week 0 and week 15), treatment replicate (15-level categorical variable; five replicates for each of the three treatments), and their interaction. The sampling fitness assessment time was considered as random variable. Pairwise comparisons of the least square means were performed between weeks within replicates and within weeks. Adjustment for multiple comparisons was based on Tukey's method. We expected that populations that underwent a (stronger) bottleneck had fewer replicates with significant fitness increase.
(iii) 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 measurements nested within replicate) for each of the three treatments. We fitted a nested ANOVA model to the population sizes with week (2-level categorical variable, week 0 and week 15), replicate (5-level categorical variable corresponding to the five replicates), and fitness assessment time (9-level factor corresponding to the multiple measurements done for each replicate; max. 3 assessment times with 3 measurements per assessment time, Table S1) as fixed effect, with measurement nested within replicate [aov(population_size ~ week + replicate/measurement]. We then extracted the mean squares and calculated for each effect what the relative proportion of the mean squares of that effect was. Because the data were unbalanced due to variation in the number of measurements done per replicate (between 3 and 9; if there was only one assessment time or three assessment times respectively, Table S1), we subsampled the data 100 times to always have one assessment time (and hence three measurements) per replicate and calculated variance proportions for each iteration. We expected that populations that underwent a (stronger) bottleneck had more variance in fitness attributable to drift relative to selection.

RESULTS

Effect of population bottlenecks on average and maximum fitness across replicates
Before selection to the novel conditions, the populations under the moderate bottleneck treatment had a significant lower fitness than the populations not exposed to a bottleneck (t-ratio = -4.779 and p-value < 0.0001) and also significantly lower than the populations under the strong bottleneck treatment (t-ratio = -7.152 and p-value <0.0001, Table 1). 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. 3). 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

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.07.463474; this version posted October 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

fitness across replicates were found among treatments (Ordered Heterogeneity test statistic = 0.584 and *p*-value = 0.5599).

Table 1: Effect of population bottleneck after selection across replicates. A) Goodness-of-fit for various distributions. B) The summary of the model using the negative binomial distribution with quadratic parameterization. C) Pairwise comparisons of the least square means both within treatments and within weeks.

A) Goodness-of-fit for distributions

Model: population_size ~ week * treatment + (1|treatment replicate) + (1|assessment time)

Distribution	Df	AIC	logLik	deviance	Chisq	Chi Df	Pr(>Chisq)
Poisson	8	278244	-139114	278228			
Negative Binomial (linear)	9	2608	-1295	2590	2.76E+05	1	<2e-16
Negative Binomial (quadratic)	9	2593	-1287	2575	1.56E+01	0	<2e-16
Tweedie	10	2594	-1287	2574	1.13E+00	1	0.29

B) Summary of fixed effects (with Negative Binomial distribution using quadratic parameterization)

	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept, i.e. week 0, no bottleneck)	7.94	0.30	26.80	<2.00E-16	***
week 15	1.82	0.24	7.66	1.83E-14	***
moderate bottleneck	-1.82	0.38	-4.78	1.76E-06	***
strong bottleneck	0.47	0.38	1.25	0.21	
week 15 : moderate bottleneck	1.41	0.32	4.48	7.57E-06	***
week 15 : strong bottleneck	-0.65	0.30	-2.16	0.03	*

C) Pairwise comparisons

within treatments between weeks

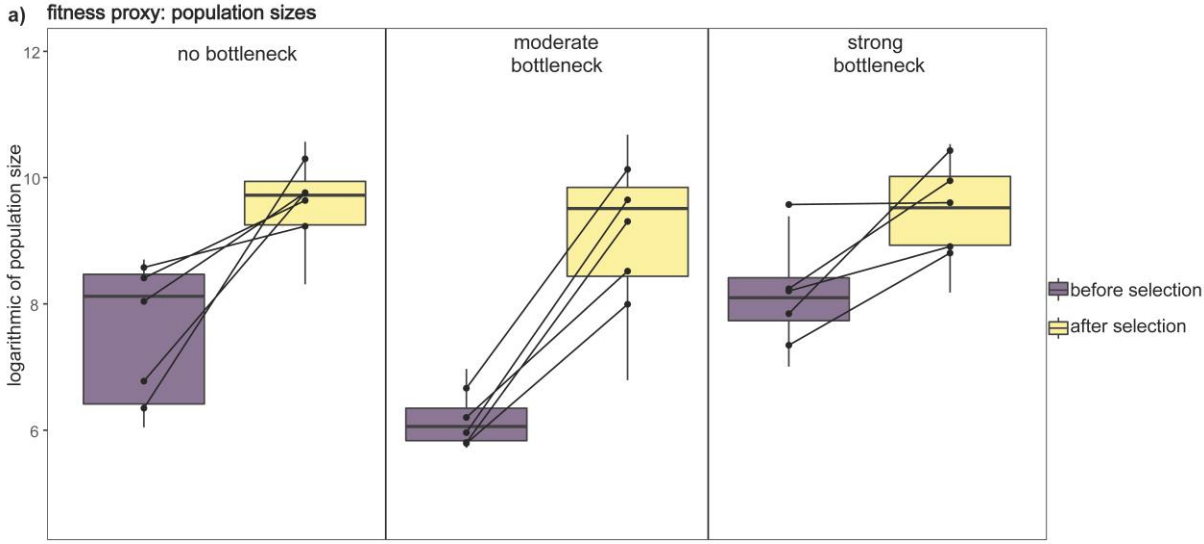
Contrast	Estimate	SE	df	t-ratio	p-value
no bottleneck (week 0 - 15)	-1.82	0.24	123	-7.66	< 0.0001
moderate bottleneck (week 0 - 15)	-3.23	0.23	123	-13.94	< 0.0001
strong bottleneck (week 0 - 15)	-1.16	0.23	123	-5.08	< 0.0001

within weeks between treatments

Contrast	Estimate	SE	df	t-ratio	p-value
week 0 (no - moderate)	1.82	0.38	123	4.78	< 0.0001
week 0 (no - strong)	-0.47	0.38	123	-1.25	0.426
week 0 (mod. - strong)	-2.29	0.32	123	-7.15	< 0.0001
week 15 (no - moderate)	0.41	0.30	123	1.34	0.376
week 15 (no - strong)	0.19	0.30	123	0.61	0.814
week 15 (mod. strong)	-0.22	0.29	123	-0.76	0.728

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.07.463474; this version posted October 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

a) fitness proxy: population sizes



b) means and medians of fitness proxy

	before selection		after selection	
	mean	median	mean	median
no bottleneck	2871	3360	17270	16660
moderate bottleneck	465	428	13892	13510
strong bottleneck	5185	3285	15848	13683

Figure 3: (a) $\log_{10}$-scaled population sizes after seven days of growth on B. megaterium before and after selection. Box-and-whisker plots show distributions across measurements (three measurements per assessment time, per replicate, between 1 and 3 assessment times per replicate). Black lines connect replicate averages (across measurements) before and after selection. **(b) means and medians of fitness proxy.** The means and medians are provided per treatment before and after selection.

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 = 5,118; Fig 4a) based on the Ordered Heterogeneity Test (r_sP_c statistic = 0.327 and p-value = 0.228). The alternative hypothesis where the bottleneck treatments did not differ from each other but obtained a lower selection response than the treatment without bottleneck was also not significant (r_sP_c statistic = 0.573 and p-value = 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 3; Fig 4a).

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.07.463474; this version posted October 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

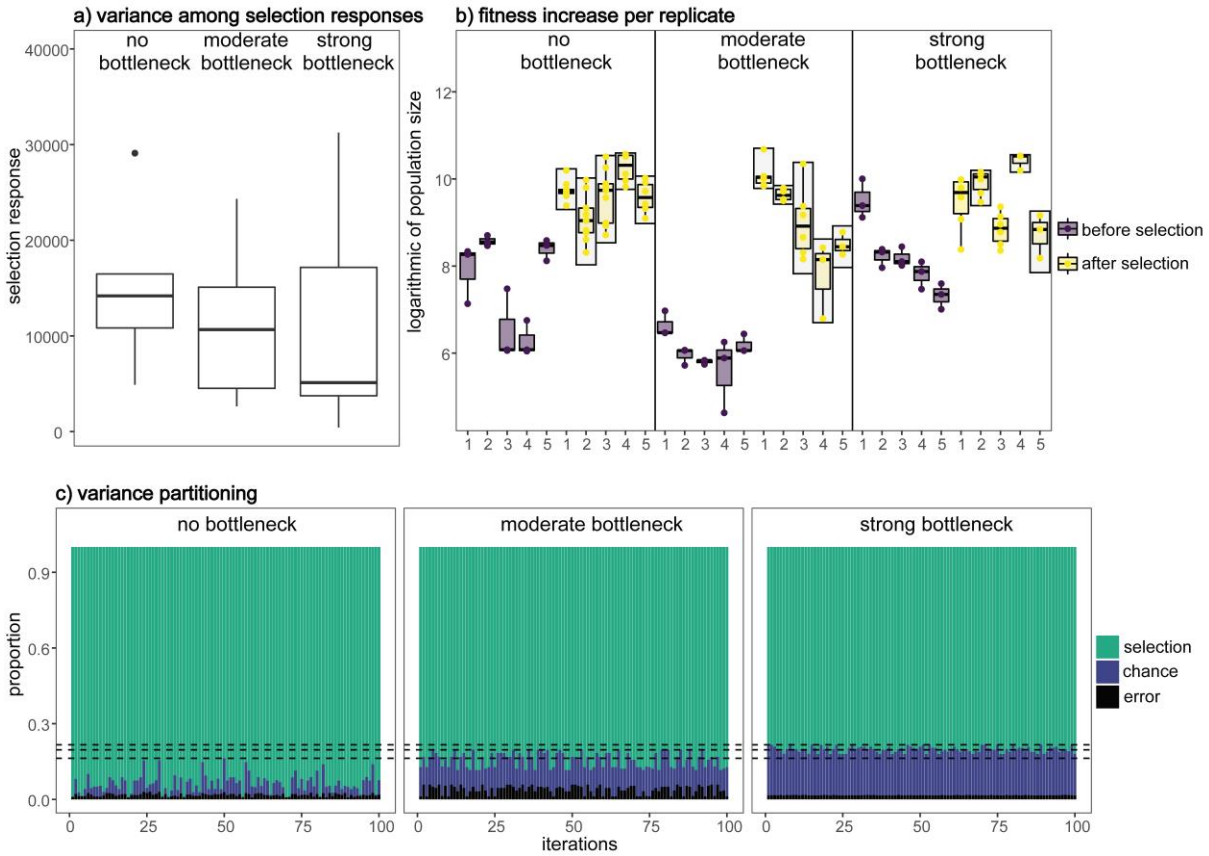


Figure 4: Effect of population bottleneck on repeatability. a) Selection response, i.e. the difference in population size after 7 days of growth on *B. megaterium* between week 0 (unadapted) populations and week 15 (putative adapted) populations. **b) Fitness increase per replicate.** Population size after 7 days of growth on *B. megaterium* by replicate (numbered with 1-5 per treatment) populations before (purple) or after 15 weeks (yellow) of experimental evolution is shown for populations derived from an ancestor that did not undergo a bottleneck or that underwent a moderate or strong bottleneck. Box-and-whisker plots show distributions of population sizes across count replicates. Grey outlines indicate a significant shift in the mean across replicates between before and after selection (GLMMs, *p*-value < 0.01, Table S2). **c) Variance partitioning between selection, chance, and error.** The data was subset 100 times on the level of technical replicates to balance data prior to calculating mean squares and to subsample the data to evaluate sampling variance in the variance proportions. For every iteration (on the *x*-axis), bars show the relative proportion of the mean squares (on the *y*- axis) that is attributable to selection (across-replicate variance between measurements before and after selection, in green), to chance effects (among-replicate variation, in blue), and to measurement error (in black). The dashed horizontal lines delineate the maximum relative proportion by chance + error for every treatment, showing that the maximum relative proportion of no bottleneck subsets overlaps with the minimum of moderate bottleneck subsets, but no such overlap exists between either of those and the strong bottleneck subsets. The total possible iterations with unique combinations of sample measurements is 108 for the no bottleneck treatment and 8 each for the moderate and strong bottleneck treatments.

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.07.463474; this version posted October 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

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 and p-value = 0.182; Fig 4a). The alternative hypothesis with only a larger variance in the strong bottleneck compared to the other two treatments was also not supported (r_sP_c statistic = 0.671 and p-value = 0.056).

A significant fitness increase was found for all five replicates in the no bottleneck treatment and the moderate bottleneck treatment, but only for three replicates in the strong bottleneck treatment (Table S2, Fig 4b; as there are five replicates within each treatment, the p-value needs to be lower than 0.05/5 = 0.01). One of the replicates that failed to increase fitness beyond the significance threshold in the strong bottleneck treatment had exceptionally high fitness on *B. megaterium* prior to the evolutionary experiment; the other replicate had low fitness both before and after 15 weeks of evolution.

We partitioned variance in the population sizes between effects from selection (variance between before and after evolution), chance (variance among replicates), and measurement error (variance among measurements nested within replicates). Because the design was unbalanced, we permutated this analysis across the different combinations of population size assays. The distributions of relative proportions of variance explained by error, chance, and selection were then compared 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 4c). Because the variance attributable to measurement error was not different among treatments, the relative variance attributable to selection was thus lower following 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. However, predictions from theory are ambiguous, especially for sexual species, and empirical data on the repeatability of evolutionary trajectories are biased towards asexual unicellular species. We therefore tested experimentally whether populations with larger effective population sizes evolve higher average and maximum fitness and along more predictable evolutionary trajectories in an obligately outcrossing strain of the nematode *C. elegans.* We found similar average and maximum fitness across three different bottleneck treatments. Fitness variation among replicates after selection was somewhat higher in bottlenecked populations compared to populations with no bottleneck, but these differences were not statistically significant. In line with an expected effect of reduced genetic diversity

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.07.463474; this version posted October 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

on the relative contribution of stochastic effects, we found that in the strongest bottleneck treatment more variance in fitness was due to chance effects and less due to selection. Thus, a population bottleneck did shift the balance between stochastic and deterministic contributions to evolution, but only marginally influenced the repeatability of fitness evolution.

The effect of a bottleneck on average and maximum fitness

We observed that nearly all replicate *C. elegans* populations, irrespective of bottleneck treatment, evolved higher relative fitness under novel conditions. Given the relatively short time of our evolutionary experiment (~15 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. A population bottleneck was therefore expected to reduce genetic variation available to selection throughout the experiment, with negative effects on the ability of a bottlenecked population to evolve higher fitness. However, we found that a population bottleneck did not significantly impact fitness evolution (as measured by the mean or maximum fitness across replicates); although, we did find that populations that underwent a strong bottleneck tended to have both slightly lower average fitness and slightly lower average selection responses. Together, these findings 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 effectively reduce the effective population size and thus left genetic diversity unaffected. Our experimental design allowed us to analyze the effect of reduced diversity independent of population (census) size, because the experiments were initialized with 500 worms that were either transferred directly from the founder (500 founders) or that were expanded from five or fifty initial founding females. Since we do not know the number of males that contributed to the 500 founders in the moderate and strong population bottlenecks, we were unable to estimate N_e. However, fitness data support an effect of the bottleneck treatment pre-selection: variation in fitness between replicates from the week 0 populations is lower after a bottleneck, compared to the no bottleneck treatment, although this was only statistically significant in the case of the moderate bottleneck. In addition, when we partitioned the variance, we found an increased contribution of chance effects 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.

An alternative explanation is that reduced genetic variation does not necessarily reduce the adaptive potential, especially when selection is strong. This hypothesis is support both by lab studies, such as in *E. coli* adapting to thermal stress where the effect of selection strongly outweighed the effects of a population bottleneck (Wein and Dagan, 2019), and by natural observations, in particular in invasive species (Schrieber and Lachmuth, 2017). Negative effects of reduced diversity on adaptive potential may be further alleviated if genomic evolution in *C. elegans* in response to a hermaphroditic, self-fertilizing mating system has

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.07.463474; this version posted October 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

resulted in reduced susceptibility to inbreeding depression. In the wild, hermaphroditic *C. elegans* experience frequent cycles of exponential population growth followed by population crashes (Frézal and Félix, 2015). There is wide-spread epistasis underlying behavioral and fitness traits in *C. elegans* (Gaertner et al., 2012; Noble et al., 2017). These epistatic interactions may have evolved as a result of adaptation to a self-fertilizing life history because epistasis is thought to reduce the effect of selfing on inbreeding depression (Abu Awad and Roze, 2020). We therefore suggest that our results fit a paradigm in which adaptive potential 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 and Lachmuth, 2017).

The effect of a bottleneck on repeatability of fitness evolution

We expected that experimental evolution in populations with reduced genetic diversity due to a bottleneck, would increase the relative contribution of drift versus selection, resulting in more heterogeneity in the evolutionary trajectories and thus more variation in the selection response. In contrast, although we observed an effect of a strong bottleneck on the relative contribution of drift to variation in fitness, we did not find differences in the repeatability of fitness evolution across bottleneck treatments. Variance among selection responses was somewhat higher in the strong bottleneck treatment compared to the other treatments, although not significantly. Two replicate populations of the strong bottleneck treatment did not significantly increase fitness during the experiment, one of which may be explained by high initial fitness, but the other not. Therefore, these results suggest a marginal effect of a bottleneck on the predictability of fitness evolution. At the same time, we saw that around 20% of the mean-squared variation in fitness was attributable to chance events, which, given only 15 generations, can be assumed to be mainly drift. Only about 10% of the mean squares was attributed to chance in the absence of a bottleneck. This means that (i) selection is an important driver of variation in fitness in our experiment regardless of the occurrence of a bottleneck, (ii) a bottleneck has a significant effect on the balance between selection and drift, (iii) the expectation that a shift in that balance translates to lower repeatability (and thus predictability) needs additional consideration.

An important insight following from our results and in line with previous empirical (Mahrt et al., 2021; Miller et al., 2011; Wein and Dagan, 2019) and theoretical findings (Szendro et al., 2013), is that constraints on evolutionary predictability may not be uniformly influenced by effective population size. In our study, we found a bottleneck (and a resulting increase in the relative contribution of drift) to have only marginal effects on repeatability of fitness evolution in outcrossing *C. elegans*. In Lachapelle and colleagues (2015), experimentally evolved vegetative algae showed more variance in a fitness proxy in populations of small effective size compared to large populations and the authors also report a strong decrease in the relative contribution of selection (from about 80% to about 40%). Similarly,

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.07.463474; this version posted October 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

experimentally evolved *E. coli* show less parallelism in the evolution of antibiotic resistance when effective population size is small compared to large (Windels et al., 2021), but parallelism in evolution of *E. coli* thermal tolerance was unaffected by population size (Wein and Dagan, 2019). Although more data, especially for sexual species, are needed, contrasting our findings with those for asexual species suggest that recombination may modulate the effects of reduced genetic diversity on evolutionary repeatability in small populations (Weissman et al., 2010). Furthermore, looking across the different studies, the trait(s) under selection also affect the role of stochasticity in limiting repeatability, as traits closely correlated with fitness are less affected by stochastic processes compared to traits that are less closely linked to fitness (Travisano et al., 1995). This may explain why parallelism in fitness may be relatively unaffected by a bottleneck in this study as well as in thermal tolerance evolution in *E. coli*, while parallelism in allele frequency changes during *E. coli* experimental evolution was greatly reduced in small versus large populations (Wein and Dagan, 2019).

In conclusion, we find that a strong population bottleneck prior to selection in an obligate outcrossing line of *C. elegans* shifts the balance between the relative contribution to fitness variation from selection and drift, similar to observations in asexual organisms. Importantly, due to our experimental setup we can exclude that this effect is due to (collinear) differences in census population size. The effects on the evolution of fitness are marginal, as only minor differences in the fitness increase over the selection period and in the repeatability of this fitness increase were observed. Our results add to a growing body of work on the role of bottleneck and effective population size variation in the repeatability of evolution that suggest that the relationships between genetic diversity, the effect of selection, and the predictability of evolutionary change are context-dependent. The similarities and variation across study systems underline the value of complementary experiments across traits that are variably correlated to fitness, across levels of biological organization, and across the tree of life.

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.07.463474; this version posted October 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

ACKNOWLEDGEMENTS

We would like to thank Juliane Teapal and the Utrecht University Large-Particle Flow Cytometry Facility (UU-LPC) for their help with the BioSorter, Suzanne Wiezer from Aquatic Ecology at NIOO-KNAW for using the Petri plate filling machine, and Sanne van der Steen. We also acknowledge Arjan de Visser, The Predicting Evolution consortium (Dries Bonte, Mirte Bosse, Steven Declerck, Marjon de Vos, Rampal S. Etienne, Steven Goossens, Martien Groenen, Paulien Hogeweg, Jan Kammenga, Ken Kraaijeveld, Martine Maan, Frederik Mortier, Ido R. Pen, Joost Riksen, Isabel Smallegange, Maurijn van der Zee, Sander van Doorn, Koen Verhoeven, Bregje Wertheim, Suzanne Wiezer, Lars E. Zandbergen), and Origins Center for helpful discussions.

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.

DATA ACCESSIBILITY

The data and R code will be made available on Figshare upon acceptance.

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.07.463474; this version posted October 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

REFERENCES

Abu Awad, D. and Roze, D. (2020). Epistasis, inbreeding depression, and the evolution of self-fertilization. *Evolution (N. Y).* **74**, 1301–1320.

Barrick, J. E., Deatherage, D. E. and Lenski, R. E. (2020). A Test of the Repeatability of Measurements of Relative Fitness in the Long-Term Evolution Experiment with Escherichia coli. In *Evolution in Action: Past, Present and Future: A Festschrift in Honor of Erik D. Goodman* (ed. Banzhaf, W.), Cheng, B. H. C.), Deb, K.), Holekamp, K. E.), Lenski, R. E.), Ofria, C.), Pennock, R. T.), Punch, W. F.), and Whittaker, D. J.), pp. 77–89. Cham: Springer International Publishing.

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

Blount, Z. D., Lenski, R. E. and Losos, J. B. (2018). Contingency and determinism in evolution: Replaying life's tape. *Science (80-.).* **655**,.

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

Brooks, M. E., Kristensen, K., van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A., Skaug, H. J., Mächler, M. and Bolker, B. M. (2017). glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *R J.* **9**, 378–400.

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

Colosimo, P. F., Hosemann, K. E., Balabhadra, S., Villarreal, G., Dickson, M., Grimwood, J., Schmutz, J., Myers, R. M., Schluter, D. and Kingsley, D. M. (2005). Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. *Science (80-.).* **307**, 1928–1933.

De Vos, M. G. J., Dawid, A., Sunderlikova, V. and Tans, S. J. (2015). Breaking evolutionary constraint with a tradeoff ratchet. *PNAS* **112**, 14906–14911.

Elmer, K. R., Fan, S., Kusche, H., Spreitzer, M. L., Kautt, A. F., Franchini, P. and Meyer, A. (2014). Parallel evolution of Nicaraguan crater lake cichlid fishes via non-parallel routes. *Nat. Commun.* **5**, 1–8.

Frézal, L. and 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. and 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. and Noguchi, K. (2020). lawstat: Tools for Biostatistics, Public Policy, and Law.

Gavrilets, S. (2004). *Fitness Landscapes and the Origin of Species (MPB-41)*. Princeton University Press.

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.

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.07.463474; this version posted October 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

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*. *Mol. Biol. Evol.* **34**, 831–842.

Gray, J. C. and Cutter, A. D. (2014). Mainstreaming Caenorhabditis elegans in experimental evolution. *Proc. R. Soc. B Biol. Sci.* **281**,.

Jain, K., Krug, J. and Park, S.-C. (2011). Evolutionary advantage of small populations on complex fitness landscapes. *Evolution (N. Y).* **65**, 1945–1955.

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

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

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

Lachapelle, J., Reid, J. and Colegrave, N. (2015). Repeatability of adaptation in experimental populations of different sizes. *Proc. R. Soc. B* **282**, 1–8.

Lässig, M., Mustonen, V. and Walczak, A. M. (2017). Predicting evolution. *Nat. Ecol. Evol.* **1**, 77.

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

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

MacNeil, L. T., Watson, E., Arda, H. E., Zhu, L. J. and Walhout, A. J. M. (2013). Diet-induced developmental acceleration independent of TOR and insulin in C. elegans. *Cell* **153**, 240–252.

Mahrt, N., Tietze, A., Künzel, S., Franzenburg, S., Barbosa, C., Jansen, G. and Schulenburg, H. (2021). Bottleneck size and selection level reproducibly impact evolution of antibiotic resistance. *Nat. Ecol. Evol.* **5**, 1233–1242.

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

Neuhäuser, M. and 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. and Sandoval, C. P. (2002). Host-plant adaptation drives the parallel evolution of reproductive isolation. *Nature* **417**, 440–443.

Poelwijk, F. J., Kiviet, D. J., Weinreich, D. M. and Tans, S. J. (2007). Empirical fitness landscapes reveal accessible evolutionary paths. *Nature* **445**, 383–386.

Rice, W. R. and Gaines, 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. and De Visser, A. J. G. M. (2008). Heterogeneous adaptive trajectories of small populations on complex fitness landscapes.

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.07.463474; this version posted October 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

PLoS One **3**, e1715.

Schrieber, K. and 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. and Avery, L. (2006). Dietary choice behavior in *Caenorhabditis elegans*. *J. Exp. Biol.* **209**, 89–102.

Stiernagle, T. (2006). Maintenance of *C. elegans*. *WormBook, ed.*

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

Travisano, M., Mongold, J. A., Bennett, A. F. and Lenski, R. E. (1995). Experimental tests of the roles of adaptation, chance, and history in evolution. *Science (80-.).* **267**, 87–91.

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

Watson, E., Yilmaz, L. S. and Walhout, A. J. M. (2015). Understanding Metabolic Regulation at a Systems Level: Metabolite Sensing, Mathematical Predictions, and Model Organisms. *Annu. Rev. Genet.* **49**, 553–575.

Wein, T. and 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. and Fisher, D. S. (2010). The rate of fitness-valley crossing in sexual populations. *Genetics* **186**, 1389–1410.

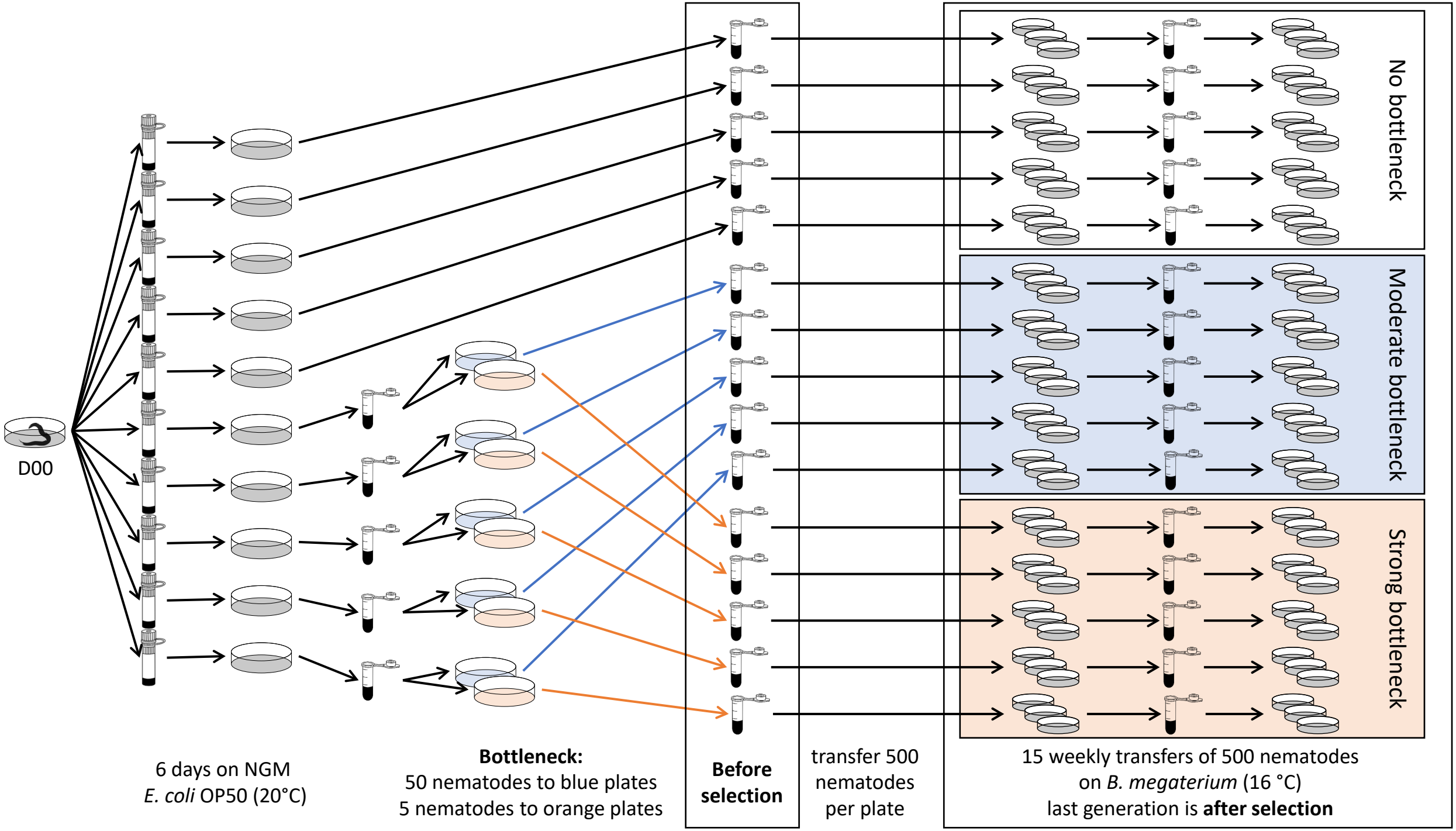
Willi, Y., Buskirk, J. Van and Hoffmann, A. A. (2006). Limits to the adaptive potential of small populations. *Annu. Rev. Ecol. Evol. Syst.* **37**, 433–458.

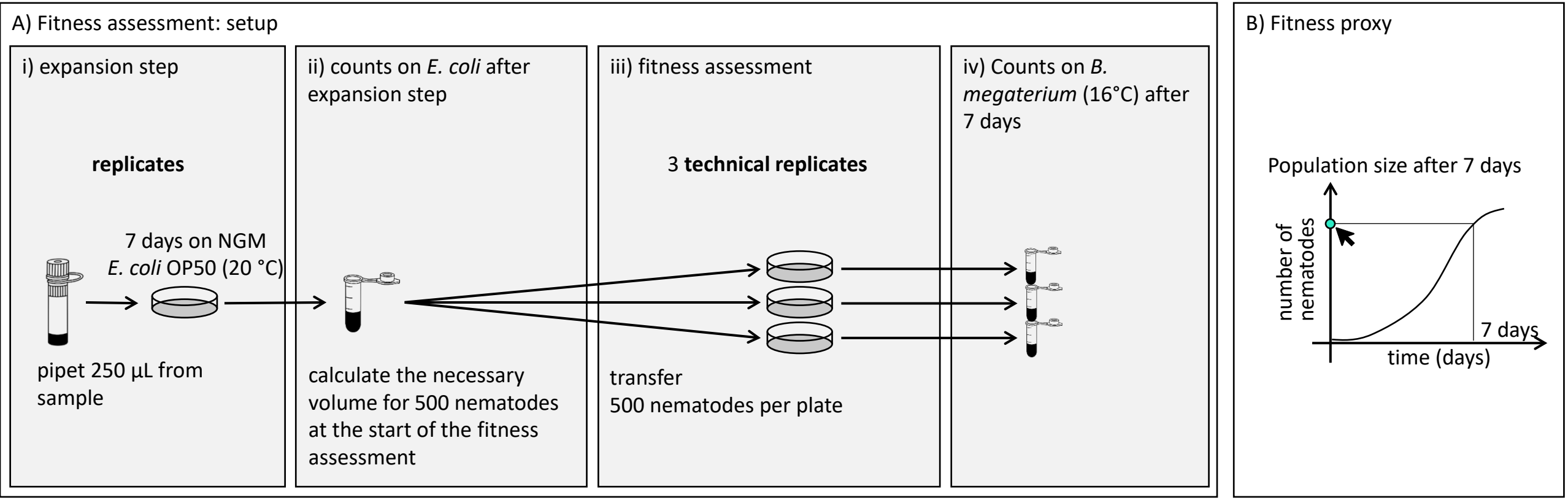
Windels, E. M., Fox, R., Yerramsetty, K., Krouse, K., Wenseleers, T., Swinnen, J., Matthay, P., Verstraete, L., Wilmaerts, D., Van den Bergh, B., et al. (2021). Population Bottlenecks Strongly Affect the Evolutionary Dynamics of Antibiotic Persistence. *Mol. Biol. Evol.* **38**, 3345–3357.

Wortel, M. T., Agashe, D., Bailey, S. F., Bank, C., Bisschop, K., Blankers, T., … and Pennings, P. S. (2021). The why, what and how of predicting evolution across biology: from disease to biotechnology to biodiversity. *EcoEvoRxiv*.

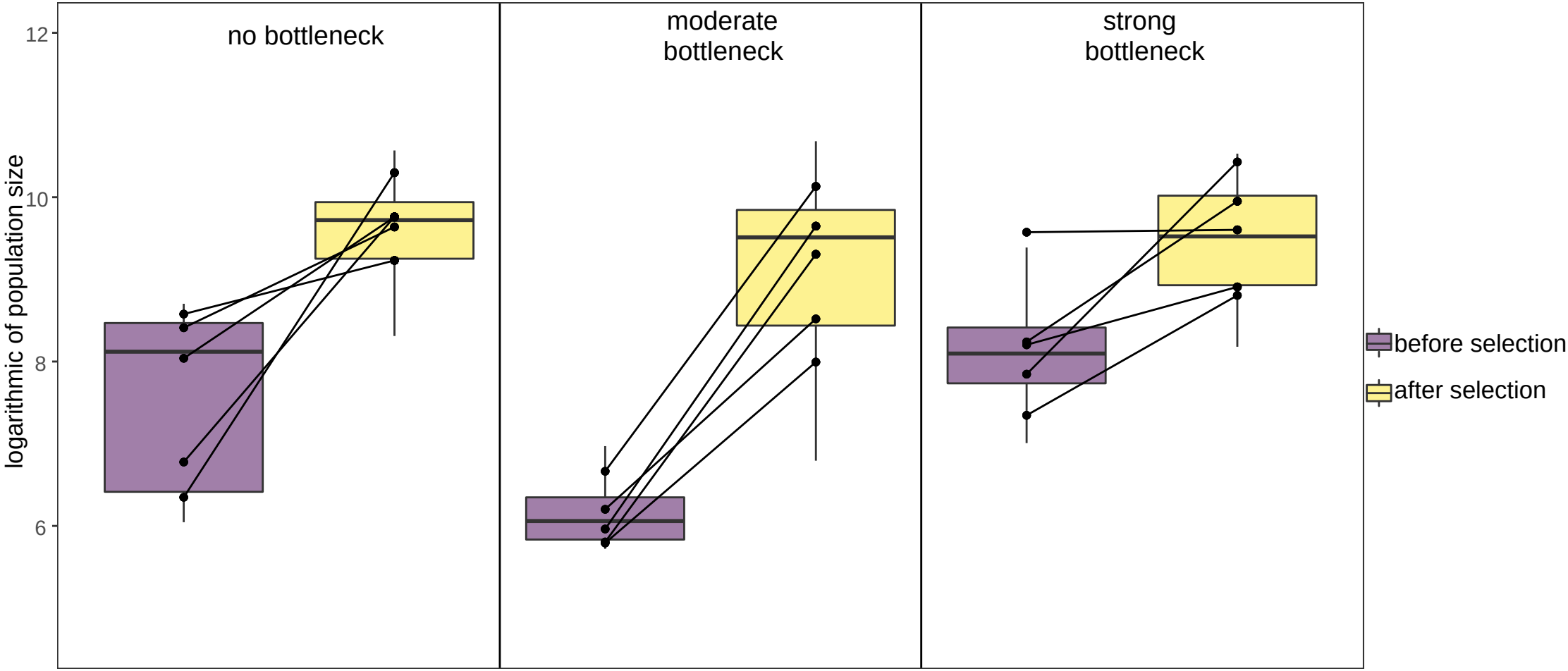
Wright, S. (1931). EVOLUTION IN MENDELIAN POPULATIONS. *Genetics* **16**, 97–159.

Wright, S. (1932). The roles of mutation, inbreeding, crossbreeding and selection in evolution. *Proc. Sixth Int. Congr. Genet.* 356–366.


D00
6 days on NGM
E. coli OP50 (20°C)
Bottleneck:
50 nematodes to blue plates
5 nematodes to orange plates
Before selection
transfer 500 nematodes per plate
No bottleneck
Moderate bottleneck
Strong bottleneck
15 weekly transfers of 500 nematodes on *B. megaterium* (16 °C)
last generation is **after selection**

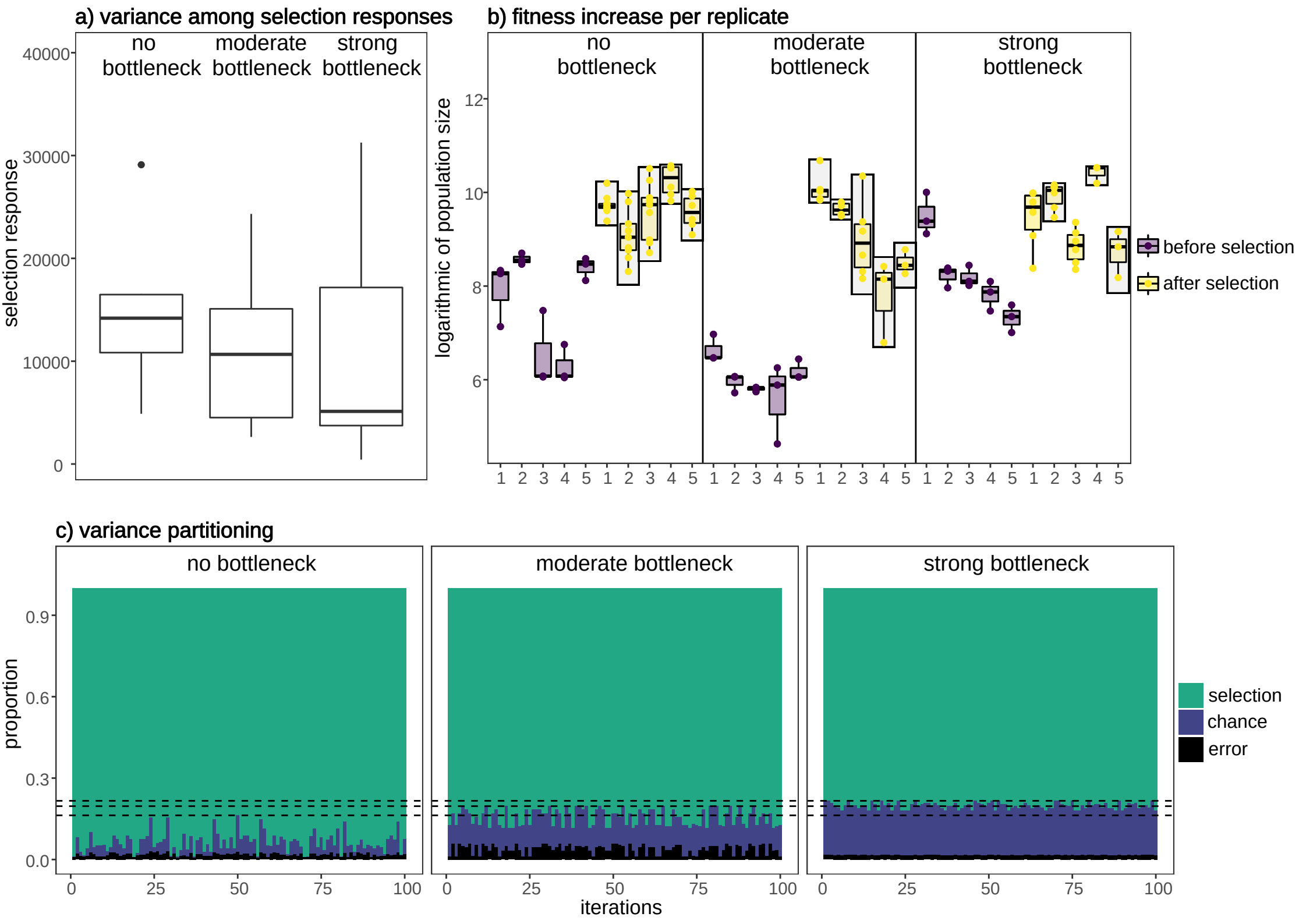

A) Fitness assessment: setup
i) expansion step
replicates
7 days on NGM
E. coli OP50 (20 °C)
pipet 250 µL from sample
ii) counts on E. coli after expansion step
calculate the necessary volume for 500 nematodes at the start of the fitness assessment
iii) fitness assessment
3 technical replicates
transfer
500 nematodes per plate
iv) Counts on B. megaterium (16°C) after 7 days
B) Fitness proxy
Population size after 7 days
number of nematodes
7 days
time (days)

a) **fitness proxy: population sizes**



b) **means and medians of fitness proxy**

	before selection		after selection	
	mean	median	mean	median
no bottleneck	2871	3360	17270	16660
moderate bottleneck	465	428	13892	13510
strong bottleneck	5185	3285	15848	13683


a) variance among selection responses
no bottleneck
moderate bottleneck
strong bottleneck
selection response
40000
30000
20000
10000
0
b) fitness increase per replicate
no bottleneck
moderate bottleneck
strong bottleneck
logarithmic of population size
12
10
8
6
1 2 3 4 5 1 2 3 4 5 1 2 3 4 5 1 2 3 4 5 1 2 3 4 5 1 2 3 4 5
before selection
after selection
c) variance partitioning
no bottleneck
moderate bottleneck
strong bottleneck
proportion
0.9
0.6
0.3
0.0
0 25 50 75 100
iterations
selection
chance
error