Experimental Evidence for the Negative Effects of Self-Fertilization on the Adaptive Potential of Populations

Highlights
- Response to selection is better in outcrossing than in selfing
- Purge of inbreeding depression does not help selfers respond better to selection
- Phenotypic variance does not always decrease in selfing, while genetic variance does

Authors
Elsa Noël, Philippe Jarne, Sylvain Glémin, Alicia MacKenzie, Adeline Segard, Violette Sarda, Patrice David

Correspondence
elsacnoel@gmail.com

In Brief
Noël et al. present results of a response-to-selection experiment on a morphological trait of Physa acuta experimental evolution lines with two levels of inbreeding depression under outcrossing versus selfing. In all cases, selfing erodes genetic variation faster than outcrossing, supporting the hypothesis that selfing is an "evolutionary dead end."
Experimental Evidence for the Negative Effects of Self-Fertilization on the Adaptive Potential of Populations

Elsa Noël,1,4* Philippe Jarne,1 Sylvain Glémin,2,3 Alicia MacKenzie,1 Adeline Segard,1 Violette Sarda,1 and Patrice David1
1Centre d’Ecologie Fonctionnelle et Evolutive (CEFE), UMR 5175, CNRS - Université de Montpellier - Université Paul Valéry Montpellier - EPEHE, 1919 Route de Mende, 34293 Montpellier Cedex 5, France
2Institut des Sciences de l’Evolution (ISEM - UMR 5554 Université de Montpellier - CNRS-IRD-EPHE), CC64, Place Eugène Bataillon, 34095 Montpellier Cedex, France
3Department of Ecology and Genetics, Evolutionary Biology Centre, Uppsala University, Norbyvägen 18D, 75236 Uppsala, Sweden
4Lead Contact
*Correspondence: elsacnoel@gmail.com
http://dx.doi.org/10.1016/j.cub.2016.11.015

SUMMARY

Self-fertilization is widely believed to be an “evolutionary dead end” [1, 2], increasing the risk of extinction [3] and the accumulation of deleterious mutations in genomes [4]. Strikingly, while the failure to adapt has always been central to the dead-end hypothesis [1, 2], there are no quantitative genetic selection experiments comparing the response to positive selection in selfing versus outcrossing populations. Here we studied the response to selection on a morphological trait in laboratory populations of a hermaphroditic, self-fertile snail under either selfing or outcrossing. We applied both treatments to two types of populations: some having undergone frequent selfing and purged a substantial fraction of their mutation load in their recent history [5], and others continuously maintained under outcrossing. Populations with a history of outcrossing respond faster to selection than those that have experienced selfing. In addition, when self-fertilization occurs during selection, the response is initially fast but then rapidly slows, while outcrossing populations maintain their response throughout the experiment. This occurs irrespective of past selfing history, suggesting that high levels of inbreeding depression, contrary to expectation [6], do not set strong limits to the response to selection under inbreeding, at least at the timescale of a few generations. More surprisingly, phenotypic variance is consistently higher under selfing, although it quickly becomes less responsive to selection. This implies an increase in non-heritable variance, hence a breakdown of developmental canalization [7] under selfing. Our findings provide the first empirical support of the short-term positive and long-term negative effects of selfing on adaptive potential.

RESULTS AND DISCUSSION

Self-fertilization has several short-term advantages, including a higher fidelity of gene transmission across generations, a more efficient exposition of alleles to selection, and reproductive assurance when pollinators or partners are scarce [8, 9]. However, it is also expected to decrease adaptive potential in the long term by increasing genetic drift and limiting recombination efficiency at the genome scale: selfing populations should display lower quantitative genetic variance than outcrossing populations of similar census size [10] and have a reduced ability both to respond to positive selection from standing variation [11] and to fix several advantageous alleles at once (selective interference [12]). The impact of selfing on adaptive responses to environmental change was thus initially the core of the “evolutionary dead end” argument.

We addressed this central question in the hermaphroditic outcrossing snail Physa acuta via the following questions: (1) Do populations regularly exposed to self-fertilization lose additive genetic variance (adaptive potential) compared to outcrossing ones? (2) Does self-fertilization during selection affect the population response to selection? (3) Does inbreeding depression strongly limit the ability to respond to selection under self-fertilization? We constructed two types of experimental evolution lines (each line being a laboratory population of N = 80 adults; two replicates per type): individuals from C (outcrossing) lines always outcrossed, whereas self-fertilization was imposed every other generation in S (frequent selfing) lines [5]. Each type was represented by two independent replicates. After around 30 generations, we performed two parallel selection experiments within each line, one under 100% selfing and the other under 100% outcrossing. In both cases, we measured 200 adults each generation and retained the 50 (1/4; Figure 1) with the highest shell width-to-length ratio (hereafter “shell roundness,” a trait known to be under selection in natural contexts; see Supplemental Experimental Procedures). In order to specify our expectations, we used individual-based simulations of experimental evolution followed by selection on a multilocus trait, mimicking our experimental conditions, accounting for deleterious mutations and inbreeding depression (see Supplemental Experimental Procedures).
Self-Fertilization Enhances the Response to Selection at First but Compromises It Later

In all eight populations (two types × two replicates × two mating treatments during six generations of selection; 12,000 individuals measured), we observed significant increases in shell roundness over generations (p < 0.001; Table 1; Figure 2A3), while unselected experimental evolution lines (1,200 individuals measured) underwent no significant changes during the same period (Table S1; Figure S1). Both C and S populations responded to selection under outcrossing (Figures 2A2 and 2A3), but the phenotypic change per generation, which reflects the additive genetic variance, was higher in C than in S populations in both the first and second halves of the experiment (“line type × generation”: p = 0.015 and p < 0.001, respectively; the dataset was broken in two halves, as the response was not linear over the entire period). Thus, 30 generations of frequent selfing substantially decreased the additive variance for quantitative traits, providing a positive answer to our first question. As no selection was applied on shell roundness over these 30 generations, this effect likely resulted from an acceleration of genetic drift. Gene diversity at microsatellite markers (7 microsatellites, 654 genotyped individuals) was reduced in S relative to C lines (0.271 versus 0.376; Figure S2) and was not significantly different from the loss in quantitative variance, with a realized heritability of 0.19 versus 0.33, −42% does not lie outside the estimation error around the neutral estimate (p = 0.51, two-sided test based on 1,000 bootstraps over microsatellite loci). Theory predicts that selfing accelerates genetic drift by increasing homozygosity and by reducing effective recombination, which allows linkage disequilibria to accumulate, causing selection at any locus to lower genetic diversity at linked loci. Both effects are expected to remain moderate under the alternate outcrossing/selfing regime of S lines, as heterozygosity never drops by more than 50% and linkage disequilibria do not accumulate significantly either (Figure S3). Accordingly, our individual-based simulations predicted only a small difference in realized heritability between S and C lines (on average 11%; Figure 2B1). The observed difference (Figure 2B) is significantly larger (42%, a value reached with probability p = 0.004 in 1,000 simulations), suggesting that selfing reduces effective size through additional processes not represented in the simulations. A possible explanation is maternal inbreeding effects whereby some inbred females lay low-quality eggs, potentially leading to the loss of entire families, hence reducing N_e. In the future, it would be interesting to investigate how the reduction of N_e under selfing is impacted by different genetic architectures of the mutation load (e.g., epistasis).

We addressed the second question by comparing the selfing and outcrossing treatments. Self-fertilization, when coupled with selection, should initially accelerate the response to selection because it increases the genetic parent-offspring resemblance, including the additive effects on traits, compared to random mating. Selfing treatment may also bear identical allelic combinations at one or several loci, so that dominance and epistatic interactions become inheritable. However this
Inbreeding depression on survival was indeed lower in S than may maintain their response to selection for a longer time. Populations with previous experience of self-fertilization, which to selection on shell roundness at the same time. In this context, selfing populations had to purge deleterious alleles and respond at each generation but selective interference remained, as experiment, extinction risks were removed by population regula-
are made homozygous and exposed to selection [18]. In our
in low population fitness and demographic risks but also
enough time to adapt [17]. High inbreeding depression results following an environmental change can be rescued by immigra-
tion, which relieves inbreeding depression, giving populations
in some groups (i.e., there may be a higher phenotypic change per unit of selection differential). In order to disentangle these
effects, we plotted both the phenotypic values against cumulative selection differentials (CSD) (Figures 2B2 and 2B3) and the CSD against selection generations (Figures 2C2 and 2C3).

Based on our simulations, we expected the differences in response to selection to stem from differences in the phenotypic response per unit of selection differential, rather than from changes in the selection differential itself (Figures 2B1 and 2C1). However, our results revealed unanticipated effects of selfing. In S lines, the phenotypic change per unit of CSD was initially larger under the selfing than under the outcrossing treatment, and was then lower, as predicted by simulations (treatment \times CSD interaction, p < 0.001 before and after generation 3; Figures 2B2 and 2B3; Table S2). However, the selection differential per generation remained consistently higher under selfing than under outcrossing (regression of CSD over generations, treatment \times generation interaction; p < 0.001 both before and after generation 3). There was no detectable change in time, as attested by linear increases in CSD with generation number (Figures 2C2 and 2C3; Table S3). The pattern of change in CSD with generations matched that of phenotypic variance, which was consistently larger in the selfing treatment than under outcrossing throughout the experiment (regression of phenotypic SD on generations: treatment effect, p < 0.001). This effect was not predicted by our simulations. The decrease in the response to selection under selfing, in spite of a higher phenotypic variance and selection differential, implies that the phenotypic variance under selfing includes a component not responsive to selection, and selection differential, implies that the phenotypic variance
predicted by our simulations. The decrease in the response to selection under selfing than in C lines (0.11 ± 0.12 and 0.50 ± 0.11, respectively; see [5]), but both the decrease in selection response under selfing and the difference in per-generation phenotypic change between outcrossing and selfing treatments (non-significant three-way interactions in Table 1) were similar in S and C lines. Our simulations did not uncover such a difference either, casting doubt on whether inbreeding depression can detectably alter the adaptive potential under selfing once demographic risks are removed. This deserves further exploration, given that our simulations predicted a smaller loss of quantitative variance in S lines than observed (see above) and may therefore have underestimated selfing impacts. All in all, the total phenotypic progress was lower in S than in C populations after six generations of selection (Figure 2A3), irrespective of the mating system. This shows that the advantage of lower inbreeding depression in S lines did not compensate for the loss of genetic variance that took place during their 30 generations of alternate selfing.

Origin of Differences between Selfing and Outcrossing: More Than Erosion of Genetic Variance

All lines experienced the same selection intensity because a constant fraction of the shell roundness distribution was selected per generation. Differences in the rate of phenotypic change may therefore arise for two non-exclusive reasons: (1) the phenotypic variance may be larger in some lines, resulting in a larger selection differential (i.e., the average difference in roundness between the upper quartile and the whole distribution), and/or (2) the fraction of variance that is genetically transmitted over generations may be larger in some groups (i.e., there may be a higher phenotypic change per unit of selection differential). In order to disentangle these effects, we plotted both the phenotypic values against cumulative selection differentials (CSD) (Figures 2B2 and 2B3) and the CSD against selection generations (Figures 2C2 and 2C3).

Our third question related to inbreeding depression, thought to be a limit to adaptation in inbred populations [6]. For example, experimental populations of Tribolium that are declining following an environmental change can be rescued by immigration, which relieves inbreeding depression, giving populations enough time to adapt [17]. High inbreeding depression results in low population fitness and demographic risks but also heightens selective interference effects as deleterious alleles are made homozygous and exposed to selection [18]. In our experiment, extinction risks were removed by population regulation at each generation but selective interference remained, as selfing populations had to purge deleterious alleles and respond to selection on shell roundness at the same time. In this context, populations with previous experience of self-fertilization, which have already purged a large part of their inbreeding depression, may maintain their response to selection for a longer time. Inbreeding depression on survival was indeed lower in S than in C lines (0.11 ± 0.12 and 0.50 ± 0.11, respectively; see [5]), but both the decrease in selection response under selfing and the difference in per-generation phenotypic change between outcrossing and selfing treatments (non-significant three-way interactions in Table 1) were similar in S and C lines. Our simulations did not uncover such a difference either, casting doubt on whether inbreeding depression can detectably alter the adaptive potential under selfing once demographic risks are removed. This deserves further exploration, given that our simulations predicted a smaller loss of quantitative variance in S lines than observed (see above) and may therefore have underestimated selfing impacts. All in all, the total phenotypic progress was lower in S than in C populations after six generations of selection (Figure 2A3), irrespective of the mating system. This shows that the advantage of lower inbreeding depression in S lines did not compensate for the loss of genetic variance that took place during their 30 generations of alternate selfing.

| Table 1. Regression of Individual Shell Roundness on Generations for Lines and Treatments |
|---|---|---|
| Effect | \(\chi^2\) | \(p\) |
| **Before Bp** | | |
| Generation | 51.53 | 7 \(\times\) 10\(^{-13}\) |
| Line \(\times\) generation | 5.95 | 0.015 |
| Treatment \(\times\) generation | 23.59 | 1.2 \(\times\) 10\(^{-6}\) |
| Line \(\times\) treatment \(\times\) generation | 2.63 | 0.104 |
| **After Bp** | | |
| Generation | 70.64 | <2 \(\times\) 10\(^{-16}\) |
| Line \(\times\) generation | 28.99 | 7.3 \(\times\) 10\(^{-8}\) |
| Treatment \(\times\) generation | 5.59 | 0.018 |
| Line \(\times\) treatment \(\times\) generation | 1.43 | 0.231 |

The individual phenotypic values were regressed on the number of selection generations, and the factors are experimental evolution line (C and S) and treatment (outcrossing and selfing). We considered all interactions. Line, treatment, and their interactions were modeled as fixed effects. Because a breakpoint (Bp) in the response to selection was detected (see Figure 2B2), the generation effect (response slope) was estimated both before and after the Bp. Replicate line within each line (two per line) and temporal periods of one month (21 periods) were modeled as random effects. The values reported are \(\chi^2\) and the probability associated with likelihood ratio tests (one degree of freedom in all tests). See also Tables S1–S3.
Although C lines respond better to selection, this increase in non-genetic components of variance with inbreeding has been classically observed and interpreted as a decrease in developmental homeostasis or "decanalization" [7, 19]. The reduced response to selection after a few generations of selfing may therefore reflect a combination of fast erosion of genetic variance and increased developmental instability.

Figure 2. Simulation Results and Results of the Selection Experiments in Physa acuta
Simulation results are shown in the left column; experimental results are shown in the middle and right columns. Data legend is beneath (C1). See Tables 1, S2, and S3 for the results of statistical tests associated with (A2), (B2), and (C2) respectively. See also Figures S1–S3.

(A1) Simulations for cumulative response to selection (in arbitrary phenotypic units) as a function of time (in number of selection generations) for lines and treatments. Because the response was not linear in all lines, we used a maximum-likelihood piecewise regression (details in Supplemental Experimental Procedures). We detected a significant breakpoint at the third generation (2.81; 95% confidence interval: 2.70–3.11), roughly the middle of our experiment. The model estimates regression slopes before and after the breakpoint, while imposing continuity at generation 3.

(A2) Regression of individual shell roundness on time (number of selection generations) for lines and treatments. A breakpoint is enforced at generation 3 (see Table S2).

(A3) Individual and mean ± SEM phenotypic deviations (phenotype minus the estimated mean at generation zero, before selection) per generation in all four conditions, presented side by side to improve readability.

(B1) Simulations for cumulative response to selection as a function of cumulative selection differential (CSD).

(B2) Regression of individual shell roundness on CSD for lines and treatments. A breakpoint is enforced at generation 3.

(B3) Mean phenotypic deviations (mean phenotype minus estimated mean at generation zero) as a function of CSD.

(C1) Simulations of CSD as a function of time (generations).

(C2) Regression of CSD on time (generations) for lines and treatments. A breakpoint is enforced at generation 3.

(C3) Actual CSD as a function of generations.
Conclusions
Some ideas in biology have become classical with surprisingly little empirical evidence, for example that selfing species are less able to cope with changes in selection regimes than outcrossing ones, as suggested by theoretical models [11, 14]. Although phylogenetic and genomic studies suggest higher extinction rates in selfing plants [3, 20] and a tendency to accumulate more deleterious mutations [21], we have little direct empirical evidence on how selfing modifies the adaptive potential at the scale of a few generations to a few tens of generations. Isogenic lines of Caenorhabditis elegans survive less well when selfing than when partially outcrossing under artificially elevated mutation and/or environmental changes [22]. However, isogenic lines do not mimic natural populations because they lack initial standing variation, relying on new mutations.

Our study clearly supports the idea that self-fertilization limits adaptive potential. Starting from the same standing genetic variance, populations switching to self-fertilization initially undergo an accelerated response to selection due to homozygosity and the recruitment of dominance and epistatic effects into the heritable variance. This effect is quickly annihilated by a fast erosion of genetic variance, as selfing populations become progressively unable to generate new genetic combinations through recombination. In addition, our study suggests that even discontinuous selfing is sufficient to decrease the additive genetic variance in a few tens of generations. These patterns are broadly consistent with the observation that wild populations of selfing species are usually less genetically variable than outcrossing ones at both neutral markers [22, 23] and phenotypic traits [18]. In nature, genetic erosion will obviously depend on parameters such as population size and selfing rate. Our lines are rather small, and selfing rates varied between 50% (periodic selfing before selection started) and 100% (during selection), but we did not place ourselves in extreme low-standing variance conditions. Indeed, the level of genetic diversity of selfing species is usually of the same order as that of our laboratory populations [24]. Moreover, the timescale at which self-fertilization limits the adaptive potential in our experiments is remarkably short: six generations of pure selfing or thirty generations of periodic selfing. It is therefore likely that recurrent selfing reduces the adaptive potential of populations in a large range of naturally occurring selfing rates and population sizes. Of course, this also depends on the genetic architecture of traits under selection: selfing in large populations is unlikely to limit the fixation of strongly beneficial alleles arising by mutation but still reduces the standing variance due to small-effect alleles and the possibility of combining several beneficial mutations.

Our study also yielded some unexpected results. First, inbreeding depression did not appear to strongly limit adaptation in selfing populations. Populations suddenly exposed to inbreeding through a reduction in size [6, 25] or an increase in the selfing rate [26] expose their recessive load to selection. The resulting genetic deaths may incur demographic risks and decrease the ability to simultaneously respond to other selection pressures [6], for example selection on increased shell roundness. However, previous experience with selfing and the associated partial purging of inbreeding depression was never an advantage in our study; moreover, our simulations did not predict that it should be. More studies are needed to investigate how specific conditions (e.g., effective population size, selfing rate, genetic architecture) affect such an advantage. A second unexpected result was the increase in non-genetic components of phenotypic variance, further impeding the response to selection. This effect, referred to as reduced genetic homeostasis, has already been observed [7] but has been overlooked as a potential cause of the low efficacy of adaptation in selfing populations and has usually been ignored in genetic models.

Our results suggest that adaptive challenges are less likely to be overcome by selfing than by outcrossing populations. This is relevant for many plants of agronomical interest, in which increased selfing rates are associated with strong selection during domestication [27]. Early cultivators may have benefited from the short-term positive effect of selfing on the response to directional selection, allowing the rapid fixation of agronomically interesting traits. However, the long-term preservation of adaptive potential requires a dose of outcrossing among lines, as in wheat dynamical management strategies [28]. Similarly, the loss of pollinators [29] may result in increased selfing rates [30] in plant populations that are at the same time facing climate change and biological invasions, possibly requiring rapid adaptation. Selfing certainly provides reproductive assurance, but our results also suggest that it increases the risk of losing genetic variation and failing to adapt when prolonged over more than a few generations.

SUPPLEMENTAL INFORMATION
Supplemental Information includes three figures, three tables, and Supplemental Experimental Procedures and can be found with this article online at http://dx.doi.org/10.1016/j.cub.2016.11.015.

AUTHOR CONTRIBUTIONS
E.N., P.J., S.G., and P.D. designed the study, interpreted the results, and wrote the manuscript. E.N., A.M., A.S., V.S., and P.D. performed the experiments and genetic analyses. E.N. and P.D. monitored the experimental evolution lines and analyzed the results.

ACKNOWLEDGMENTS
We thank B. Pelissié, T. Bataillon, H. Fréville, S. Maurice, and J. Ronfort for discussions; two anonymous reviewers for comments; and the SMGE (CEFE) and Environmental Genomics platform (LabEx CeMEB) for support with the molecular analyses. This work was a part of E.N.’s PhD project and was funded by the Agence Nationale de la Recherche (ESHAP, ANR-12-BSV7-0015, and AFFAIRS, ANR-12-ADAP-005 to P.D.) and by CNRS.

Received: July 27, 2016
Revised: October 12, 2016
Accepted: November 3, 2016
Published: December 29, 2016

REFERENCES


Supplemental Information

Experimental Evidence for the Negative Effects of Self-Fertilization on the Adaptive Potential of Populations

Elsa Noël, Philippe Jarne, Sylvain Glémin, Alicia MacKenzie, Adeline Segard, Violette Sarda, and Patrice David
Figure S1. Mean phenotypic deviations in the experimental treatments and in the control. Results are reported per line (C: black; S: red), replicate per line (C1: square; C2: triangle; S1: circle; S2: diamond) and treatment (outcrossing: continuous lines; selfing: dotted lines; control: dash-dotted lines). Mean phenotypic deviations were obtained by subtracting a reference value representing the mean of unselected individuals. This reference was the mean of the initial generation in the experimental treatments and the phenotypic mean over all generations in the control (no significant temporal change was detected; Table S5), and the vertical bars are standard errors. Artificial selection began between the second and third point for each replicate, as individuals were also measured at the generation preceding the one in which artificial selection was first enforced. Generation time was not constant, mostly because self-fertilization induces a delay in egg-laying, so selfing lines tended to have longer generation times. Here, we illustrated this by plotting results on an actual time axis (in weeks since the beginning of the experiment) – corresponding plots made on a per-generation or per-selection differential basis are in the main text, Figure 2. See also Figure 2.
Figure S2. Genetic diversity before the initiation of selection (G0) and after six generations of selection (G6) per line, treatment and replicate. Shaded bars are gene diversity (expected heterozygosity under Hardy-Weinberg equilibrium) and open bars denote observed heterozygosity. C/Out: black continuous; C/Self: black dotted; S/Out: red continuous; S/Self: red dotted. 1 and 2 refer to the two replicates per line.
Comment: the observed heterozygosity is initially lower in selfing than in outcrossing lines, and, as expected, even lower after six generations of selfing. See also Figure 2.
Figure S3. Results of simulations of the evolution of linkage disequilibrium ($r^2$) as a function of the recombination rate under various mating regimes. Red = 100% selfing, blue= 100% outcrossing, purple = 50% selfing, black= one generation of 100% selfing alternating with one generation of 100% outcrossing, as in our experimental lines.

Method: we considered two bi-allelic loci with an initial allelic frequency of 0.5 and no linkage disequilibrium (LD) in an initially panmictic population of effective size 400 and a given recombination rate between the two loci. After 400 generations of simulations (burn-in period), the population was assumed to be at equilibrium. A symmetric mutation at rate 0.0001 (per generation) was added at both loci to ensure enough polymorphism at the end of the burn-in period. Starting from these initial populations, the four mating regimes were applied during 30 generations without mutations and LD, conditioned on both loci being polymorphic, was computed. 10000 replicates were run for each recombination rate and mating regime (error bars around mean LD in the Figure). The two loci were polymorphic at the end of 5500 to 6500 runs depending on the conditions. The simulation code was written in R (https://www.R-project.org). Scripts are available upon request from the authors. See also Figure 2.
Table S1. Regression of shell roundness on the number of generations for control (unselected) treatment of experimental lines. The shell roundness was regressed on generations separately for the four replicates of the experimental evolution line (C and S) (a first model including the four replicates returned a significant interaction between replicate and generation, hence the separate regressions). Block was modeled as a random effect. Legend as in Figure S1. See Figure S1 for a representation of data. See also Table 1 and Figure S1.

<table>
<thead>
<tr>
<th>Replicate</th>
<th>Effect</th>
<th>Chi-2</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>Generation</td>
<td>0.75</td>
<td>0.389</td>
</tr>
<tr>
<td>C2</td>
<td>Generation</td>
<td>1.75</td>
<td>0.186</td>
</tr>
<tr>
<td>S1</td>
<td>Generation</td>
<td>1.90</td>
<td>0.168</td>
</tr>
<tr>
<td>S2</td>
<td>Generation</td>
<td>4.32</td>
<td>0.073</td>
</tr>
</tbody>
</table>

Table S2. Regression of individual shell roundness on cumulative selection differential for lines and treatments. The phenotypic values were regressed on cumulative selection differential (CSD), and the factors were experimental evolution line (C and S) and treatment (Out and Self). The breakpoint was not estimated, but enforced at generation 3 (see Fig. 1B) for the sake of comparison with results in Table 1. Legends as in Table 1. The number of degrees of freedom (df) was one in all tests. Interaction Replicate* CSD and Block were modelled as random effects. See also Table 1 and Figure 2.

<table>
<thead>
<tr>
<th>Effect</th>
<th>Chi-2</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before Bp</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CSD</td>
<td>35.51</td>
<td>2 10^-9</td>
</tr>
<tr>
<td>Line * CSD</td>
<td>16.91</td>
<td>3.9 10^-5</td>
</tr>
<tr>
<td>Treatment * CSD</td>
<td>11.75</td>
<td>6 10^-4</td>
</tr>
<tr>
<td>Line * treatment * CSD</td>
<td>7.21</td>
<td>0.007</td>
</tr>
<tr>
<td>After Bp</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CSD</td>
<td>0.08</td>
<td>0.773</td>
</tr>
<tr>
<td>Treatment * CSD</td>
<td>17.74</td>
<td>1.5 10^-5</td>
</tr>
<tr>
<td>Line * treatment * CSD</td>
<td>2.99</td>
<td>0.084</td>
</tr>
</tbody>
</table>

Table S3. Regression of cumulative selection differential on time (generations) for lines and treatments. The cumulative selection differential (CSD) was regressed on the number of selection generations, and the factors were experimental evolution line (C and S) and treatment (Out and Self). The breakpoint was not estimated, but enforced at generation 3 (see Fig. 1C) as in Table S2. Legend as in Tables 1 and S2. See also Table 1 and Figure 2.

<table>
<thead>
<tr>
<th>Effect</th>
<th>Chi-2</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before Bp</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Generation</td>
<td>987.7</td>
<td>&lt; 2 10^-16</td>
</tr>
<tr>
<td>Line * generation</td>
<td>3.11</td>
<td>0.078</td>
</tr>
<tr>
<td>Treatment * generation</td>
<td>36.08</td>
<td>1.9 10^-9</td>
</tr>
<tr>
<td>Line * treatment * generation</td>
<td>7.82</td>
<td>0.005</td>
</tr>
<tr>
<td>After Bp</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Generation</td>
<td>2229.1</td>
<td>&lt; 2 10^-16</td>
</tr>
<tr>
<td>Line * generation</td>
<td>0.25</td>
<td>0.615</td>
</tr>
<tr>
<td>Treatment * generation</td>
<td>46.7</td>
<td>8.4 10^-12</td>
</tr>
<tr>
<td>Line * treatment * generation</td>
<td>2.46</td>
<td>0.117</td>
</tr>
</tbody>
</table>
Supplemental Experimental Procedures

Lines Maintenance. The creation of the experimental evolution lines of *P. acuta* used here has already been fully described [S1]. Briefly, individuals were collected in 10 sites near Montpellier in 2007 from which a highly diverse ancestral population was derived through two generations of intermixing. Two types of experimental evolution lines were then derived at population size of 80 adults per generation (Figure 1). Each type included two replicates. Individuals from the C lines were mass-mated which results in near 100% outcrossing in this species [S2]. The absence of heterozygote deficiencies at microsatellite data (Figure S3) is a further indication of outcrossing. Individuals from the S lines were constrained to self-fertilize every other generation, followed by a generation of pair-mating with a single partner. This regime was maintained for ca. 30 generations. The strong inbreeding depression for juvenile survival observed in natural populations was largely purged in the S lines (0.1 vs. 0.5 in C lines [S1]), but some depression was still detected on adult fecundity (0.2). The selection experiment was initiated by extracting 200 outcrossed adult virgin snails from each line (this occurred at generations 29, 31, 35 and 34 of experimental evolution for lines S1, S2, C1 and C2 respectively) and split them into two groups, 100 for the Out treatment and 100 for the Self treatment. In the Out treatment, the 100 individuals were mass-mated and their offspring were collected to enter the selection protocol as the “generation zero” (G0). In the Self treatment, the 100 individuals underwent a generation of self-fertilization without selection to generate the G0 of the selection protocol.

Response to Selection Assay. The response to selection was evaluated on the ratio of width/length (W/L), hereafter shell roundness, of individual shells. Shell morphology has a strong functional significance in mollusks as it is known to be involved in evolutionary response to predation (see [S3,S4] for general references). Experimental work in *Physa acuta* has shown that high shell roundness protects against fish predators while elongated shells deter crayfish predation [S3]; In the field, populations exposed to fish predation and wave action show on average rounder shells [S5–S7]. This largely results from well-documented phenotypic plasticity in response to wave detection or predator chemical cues [S5–S7]. In our experiment such stimuli were absent and we checked in a pilot parent-offspring study that shell roundness was heritable. Shell roundness is unlikely to be inadvertently selected under our laboratory conditions, neither to be closely associated to fitness. We opted for selection by truncation in order to get rapid responses and to detect potential differences between lines and treatments. Selection was performed over six successive generations (21 months) under four different conditions: C lines maintained under a selfing treatment (C/Self), C lines and outcrossing (C/Out), S lines and selfing (S/Self), and S lines and outcrossing (S/Out). As there were two replicates per type of line (C and S) we raised 8 selected populations in total. At each generation, individuals were kept isolated and virgin until they were measured. The 25% of individuals exhibiting the largest W/L ratio were retained as parents for the next generation. Selected individuals either remained isolated until they reproduced or were mass-mated, depending on treatment. The number of reproductive parents was 50 drawn out of 200 individuals. Up to 300 individuals were sometimes measured in the selfing treatment to compensate for mortality due to inbreeding depression. During the same time the C and S lines were also maintained without selection under their usual mating regime; these individuals were measured at each generation and served as unselected controls.

Data Analysis. The response to selection was evaluated using a linear mixed model in which the individual phenotypic value was modeled as a function of the covariate “number of generations of selection”, and of categorical factors “experimental evolution line type” (S vs. C) and “experimental treatment” (selfing vs. outcrossing). Generation, line type and treatment were modeled as fixed effects including all two- and three-way interactions. We also included line replicate (two per line type) as additional random effects, as well as a temporal block random effect corresponding to one-month periods starting from the beginning of the selection. As we were interested in temporal changes in the response to selection over the course of generations, we used a piecewise regression [S8] that allows for a single change in the slope value at some point (breakpoint) along the abscissa axis – the best breakpoint (Bp) being estimated by maximum likelihood. While a piecewise-linear function is only an approximation of how selection response might change over generations, we preferred it over a quadratic function or non-linear regressions for several reasons: first linear slopes have a more straightforward biological interpretation, especially of interaction terms, than quadratic coefficients. Second, remaining within the framework of linear mixed models allowed incorporating the random terms associated with replicates and blocks. Third, quadratic models necessarily predict ever-increasing or ever-decreasing responses – an unrealistic property since responses to selection, although they might become null, can never become negative (i.e. in a direction opposite to selection).

As the response to selection differed before and after the Bp in the piecewise regression, we considered independently the interactions among the fixed factors before and after this point. The Bp value was estimated together with the regression model using a maximum-likelihood procedure coded in R (see http://stats.stackexchange.com/questions/19772/estimating-the-break-point-in-a-broken-stick-piecewise-linear-
model-with-random-effects-in-R). At each generation, the selection differential was estimated as the difference between the mean W/L in selected individuals and in all individuals. The regression analysis was repeated using the cumulative selection differential (CSD) instead of generations, and we also regressed the CSD on the number of generations of selection. In both analyses, we introduced a Bp at generation 3, based on the previous regression (the Bp was estimated at generation 2.81).

We also used a regression analysis to estimate the average realized heritability over the whole period (6 generations). In the outcrossing treatment, the realized heritability \((h^2)\) can indeed be derived as the slope of the regression of phenotypic change on CSD. In the selfing treatment, this slope represents broad-sense heritability \((H^2)\), which differs from narrow-sense \(h^2\) as it includes a fraction of dominance and interaction variances. Nevertheless, irrespective of the underlying genetics, this slope remains indicative of the average efficiency of selection at generating a phenotypic response in the treatments.

We also monitored the evolution of the phenotypic variance (more exactly its square root, phenotypic standard deviation) across generations. It was analyzed using a regression model similar to the one presented above, with generation, line and treatment as fixed factors, and line replicate as a random factor. A Bp was enforced at generation 3.

All comparisons were conducted using chi-square likelihood-ratio tests computed using restricted maximum likelihood for random terms and maximum likelihood for fixed effects. All analyses were performed with R version 3.0.0 packages lme4 ([S9]).

**Microsatellite Variation.** The analysis was done as described in Sourrouille et al. [S10] and Escobar et al. [S11]. DNA was extracted from around 40 individuals per line, replicate and treatment at the G0 and G6 generations respectively. Variation was screened at seven polymorphic loci which displayed highly repeatable patterns (AF108762, AF10874, Pac1, Pac2, Pasu1-2, Pasu1-9, Pasu1-11). Gene diversity and the observed heterozygosity were estimated using Genetix 4.2 [S12].

**Simulation programs.** We performed individual-based simulations reproducing the essential features of our experiment, including an initial phase of experimental evolution with S and C lines, followed by a phase of selection on a phenotypic trait (shell roundness in the experiment) under two treatments (Self and Out). For simplification, the genome was modelled as a single chromosome of length 10 Morgans (10 recombination events per generation on average, with a Poisson distribution). Individuals harboured two chromosomes (inherited from their mother and father respectively), each of which could carry three types of loci affecting: (i) the selected trait (shell roundness), (ii) juvenile survival, and (iii) adult fecundity – the latter two types generate inbreeding depression.

We set the number of loci affecting the trait to 50, and the initial heritability in the founding population to 0.3; each locus had an allele with positive effect on the phenotype \((+a)\) and an allele with effect 0. The allele frequencies were chosen so that gene diversities were uniformly distributed between 0 and 0.5 across loci, and the most common allele was either the \(+a\), or the 0 allele at random. The value of \(a\) was adjusted to get an additive genetic variance of 0.3 and environmental values were drawn in a Gaussian distribution with variance 0.7. We did not simulate mutations for the quantitative trait as our focus was on response to short-term selection based on standing variance.

Loci affecting juvenile and adult fitness had a wild-type and a mutant allele, with selection coefficient \(s\) and dominance coefficient \(h\). We set \(s = 0.8\) and \(h = 0.05\) for alleles affecting early survival (mimicking recessive semi-lethals) and \(s = 0.05\) and \(h = 0.2\) for alleles affecting reproduction (small-effect loci). We started with an inbreeding depression of 0.5 at the juvenile survival phase, and 0.3 at the adult phase, based on previous observations in natural populations of our model system, the snail *Physa acuta*. We adjusted genomic mutation rates so that the resulting inbreeding depression matches the required values at mutation-selection equilibrium in a large population. The number of loci was not specified and each new mutation had a new position in the chromosome, drawn at random. Individuals initially carried a number of mutations drawn from a Gaussian distribution (mean number of mutations determined by mutation-selection equilibrium) and all mutations were unique – assuming that each of them is rare.

Generations started with a pool of N parents (N=50 during the experimental evolution phase and N=50 during the selection phase, as in our experiment). We generated offspring by drawing one (under selfing) or two (under outcrossing) parents with probabilities proportional to their fecundity. Each parent produced a gamete through recombination of its two constitutive gametes and the offspring was made by uniting the gametes produced by the two parents. New mutations were randomly added for early- and late-acting fitness genes. The early survival of offspring was then computed from its genotype at the relevant loci and the individual was retained with probability proportional to this survival. We repeated this procedure until we got 50 (experimental evolution phase) or 200 (selection phase) individuals. We then computed fecundity and trait of all individuals (the latter with an environmental variance of 0.7). In the selection phase, we retained the 50 largest phenotypes,
while we retained all individuals in the experimental evolution phase. The pool of parents was then ready to start a new generation. We recorded the phenotypic distributions before and after selection at each generation (all are based on a mean of 100 independent simulations of each line type and treatment). They were used to compute the response to selection (in phenotypic units) and the cumulative selection differential.

Simulations were implemented using Mathematica (Wolfram Research, Inc., Mathematica, Version 7.0, Champaign, IL, 2008). Scripts are available upon request from the authors.

Supplemental references


