Experimental demonstration of the impact of hard and soft selection regimes on polymorphism maintenance in spatially heterogeneous environments

Romain Gallet, Rémy Froissart, Virginie Ravigné

To cite this version:

HAL Id: hal-02623091
https://hal.inrae.fr/hal-02623091
Submitted on 26 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.
Experimental demonstration of the impact of hard and soft selection regimes on polymorphism maintenance in spatially heterogeneous environments

Running title (38 characters): Experimental test of hard and soft selection

Authors: Romain Gallet\textsuperscript{1,*}, Rémy Froissart\textsuperscript{1-2,*}, Virginie Ravigné\textsuperscript{3}

\textsuperscript{*} equally contributed to the work

Affiliations:

\textsuperscript{1} INRA, UMR 385 BGPI, Cirad TA A-54/K Campus International de Baillarguet 34398 Montpellier Cedex 5, France

\textsuperscript{2} CNRS, IRD, Université de Montpellier, UMR 5290 MIVEGEC, F-34090 Montpellier, France

\textsuperscript{3} CIRAD, UMR PVBMT, F-97410 Saint-Pierre, France

*Corresponding author: rgallet@gmail.com

Author contributions: RF had the original idea of the study; RF, RG and VR designed the experiments; RF and RG carried out the experiments; RG analyzed experimental results; VR provided theoretical predictions; RF, RG and VR wrote the paper.

Acknowledgements: We are grateful to R. E. Lenski for providing \textit{E. coli} strains. We thank C. Duperray (IRMB – Montpellier) and the Montpellier RIO Imaging platform for hosting the flow cytometry measurements, F. Débarre and B. Facon for helpful comments on the manuscript, and

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/evo.13513.

This article is protected by copyright. All rights reserved.
C. Prator for careful English editing. This study received financial support by the French Agropolis Fondation (Labex Agro-Montpellier, BIOFIS Project Number 1001-001 and E-SPACE project number 1504-004), the European Union (ERDF), and the ‘Conseil Régional de La Réunion’.

Data accessibility: The data supporting our results have been archived on the DRYAD repository website (DOI: doi:10.5061/dryad.261553q)

Conflict of Interest Statement: The authors declare no conflict of interest.

Abstract

Predicting and managing contemporary adaption requires a proper understanding of the determinants of genetic variation. Spatial heterogeneity of the environment may stably maintain polymorphism when habitat contribution to the next generation can be considered independent of the degree of adaptation of local populations within habitats (i.e., under soft selection). In contrast, when habitats contribute proportionally to the mean fitness of the populations they host (hard selection), polymorphism is not expected to be maintained by selection. Although mathematically established decades ago, this prediction had never been demonstrated experimentally. Here we provide an experimental test in which polymorphic populations of Escherichia coli growing in heterogeneous habitats were exposed to hard and soft selection regimes. As predicted by theory, polymorphism was preserved longer under soft selection. Complementary tests established that soft selection slowed fixation processes and could even protect polymorphism in the long term by providing a systematic advantage to rare genotypes.

Keywords

experimental evolution, polymorphism, coexistence, density regulation, frequency dependence, bacteria

Introduction

Genetic variation is the fuel of evolution. Understanding the ultimate forces that shape the amount of genetic variation within populations is therefore a central issue of evolutionary biology. Beyond its fundamental interest, this topic is also crucial for a number of applied issues where evolutionary potential matters. In conservation biology for instance, preserving the adaptive potential of endangered species is now a primary goal of management policies (Crandall et al. 2000). Similarly, as pathogen evolution regularly ruins management attempts (e.g., antibiotic resistance, plant resistance breakdown), managing pathogen polymorphism is becoming a growing concern (Vale 2013).
The spatial heterogeneity in selection pressures among the different habitats composing an environment has early been proposed as one of the mechanisms maintaining genetic variation in *natura*. The idea sounds intuitive, and yet it has been the subject of intense debate. Levene (1953) proposed a simple population genetic model showing conditions under which a stable local adaptation polymorphism could be maintained in a panmictic population. Dempster (1955) soon challenged this view with a very similar model, in which no polymorphism could ever be stably maintained. These two models have widely been recognized as prototypical cases of two different types of selection regimes: soft and hard selection, respectively. Soft selection appeared far more prone to polymorphism maintenance than hard selection (Christiansen 1975; Karlin and Campbell 1981).

Both authors assumed that a panmictic population inhabits a heterogeneous environment composed of several discrete habitats. At each generation, individuals disperse into a habitat and face selection: they reproduce or survive differentially depending on their local adaptation genotype (i.e., local selection is modelled as density-independent and genotype-dependent survival). In Levene’s model, density regulation occurs at the local scale, *i.e.* within habitats. The population grows until it reaches the carrying capacity of the habitat and only a fixed number of individuals will survive independently of their genotype (i.e., density regulation is modelled as density-dependent and genotype-independent survival). Surviving individuals from all habitats mix, reproduce panmictically, and the life cycle starts again. In Levene’s model, the contribution of habitats to the next generation is fixed (i.e., a given number of individuals will reproduce). Consequently, individual fitness, which is the contribution of individuals to the next generation, depends on the fitness of all other individuals sharing the same habitat. The fitness of individuals is thus relative to the maximal fitness within the habitat. In contrast, in Dempster’s model density regulation happens at the scale of the whole population, after global mixing, and before redistribution into habitats. Due to this assumption, the contribution of each habitat to the next generation is not fixed by some space or resource constraints at the local scale as in Levene’s model. On the contrary, Dempster shows that individuals contribute to the next generation in proportion to their absolute fitness (i.e., their reproductive success is not affected by other individuals sharing the same habitat), so that the genotype adapted to the most productive habitat (i.e., the most frequent or with the highest carrying capacity) invades and no polymorphism can be maintained in the long run. The assumption of a fixed vs. variable habitat contribution to the next generation later appeared so consistently important for the maintenance of stable polymorphism that it is its now what is referred to as ‘soft selection’ and ‘hard selection’ respectively (a definition we hereafter use, Christiansen 1975; Débarre and Gandon 2011).
Nurtured by this initial debate, numerous theoretical studies have attempted to identify the kind of life cycles and ecological settings that promote the maintenance of stable polymorphisms. Models have explored how spatial heterogeneity enables polymorphism maintenance in a wide array of settings. Usually one dimension of the niche such as temperature, nutrients, or a biotic feature varies within the modelled environment. It may vary gradually, as in niche breadth models (Kirkpatrick and Barton 1997; Case and Taper 2000; Mizera and Meszéna 2003), or in a discrete manner in the form of a set of habitats. In most models with discrete habitats, space is implicitly modelled. The earliest theoretical results stem from models with two patches, each corresponding to one habitat and connected by either full migration (Levene 1953; Karlin and Campbell 1981; Hedrick 1990; van Tienderen 1997; Day 2000; Doebeli and Dieckmann 2000; fine-grained environments sensu Levins 1968), limited migration (Maynard-Smith 1966; Brown and Pavlovic 1992; Meszéna et al. 1997) or biased migration (i.e. with habitat choice, de Meeûs et al. 1993; Ravigné et al. 2004; Egas et al. 2004). Now, more recent models include more complex settings. The modelled (meta)population may occupy a patchy environment, where the patches are connected by some form of migration and where each patch is characterized as belonging to one of two (rarely more) contrasted habitats (Parvinen and Egas 2004; Nurmi and Parvinen 2008). Finally the case of two discrete habitats juxtaposed in a continuous space has been explicitly modelled (Débarre and Gandon 2010; Débarre and Lenormand 2011).

After decades of study, it appears that there is no simple relationship between spatial heterogeneity and stable maintenance of local adaptation polymorphisms (see Kassen 2002; Ravigné et al. 2009; Massol 2013; Vale 2013 for reviews). According to theoretical studies, whether selection leads to the stable maintenance of diversity depends on the interaction between four factors: (1) the existence of local adaptation trade-offs (i.e. negative genetic correlations in fitness across different habitats, Levins 1962), (2) the frequency and productivity of the different habitats in the environment (Levene 1953), (3) the amount of gene flow between habitats (Maynard Smith 1966; Christiansen 1975; Débarre and Gandon 2011), and (4) the way habitats contribute to the next generation (i.e., fixed vs. variable habitat contribution, Levene 1953; Dempster 1955; Christiansen 1975; Ravigné et al. 2004).

Like all theoretical work, hard and soft selection models are per se, very schematic and some intricate interactions may have been overlooked. For instance, while natural populations may have complex population dynamics, most models assume that dispersal, selection, density regulation and mixing are distinct stages, or with very few exceptions (Egas et al. 2004), a ceiling type of population regulation (i.e., only a fixed number of individuals survive the regulation stage regardless of their genotype). In reality, migration may not be concentrated at one stage in the life cycle. Besides, selection and density regulation may occur simultaneously. Carrying capacity may vary over time either in relation to local adaptation or because of extrinsic factors. The regulation of population density, instead of resulting from contest competition (i.e, the survival of a few, Nicholson 1954) could take the form of scramble competition, where all individuals suffer from degraded demographic parameters (i.e., the suffering of all, Nicholson 1954). Some populations, as observed in host parasite systems, may...
even present cyclic population dynamics. All of these factors may have significant effects in the real world, and therefore limit the relevance of predictions made by hard and soft selection models. Verifying the validity of these predictions inherently calls for experimental testing.

Despite a vast consensus among theoreticians over the importance of the type of selection regime for polymorphism maintenance in heterogeneous environments, the concepts of hard and soft selection generally remain overlooked in the empirical literature (Roff 1993; Vale 2013). Previous studies tested the effect of spatial heterogeneity on genetic variability (reviewed in Rainey et al. 2000; Kassen 2002; see also Jasmin and Kassen 2007), most of them concluding that populations confronted with a spatially heterogeneous environment are more variable than those exposed to homogeneous environments. Yet, these experiments did not control for the selection regime imposed by serial passages and experimentally applied hard selection (except García-Dorado et al. 1991; Bell and Reboud 1997). The higher variability observed under these heterogeneous treatments resulted from transient polymorphism being less efficiently removed from heterogeneous environments than from homogeneous environments. A possible exception was the experiment by Bell and Reboud (1997) in which, despite no experimentally-imposed density regulation, local density regulation was suspected to have occurred and to have promoted higher genetic variance in heterogeneous environments as compared to homogeneous ones. One study explicitly imposed hard and soft selection regimes on a mixture of strains of the unicellular algae *Chlamydomonas reinhardtii* maintained in a heterogeneous environment for 50 generations without sexual reproduction (Bell 1997). Contrary to theoretical predictions, genetic variation remained similar regardless of the type of density regulation. This unexpected result was interpreted as a consequence of the specific nature of the environmental heterogeneity – habitats were composed of different mixtures of nutrients – that did not impose a trade-off in local adaptation (Bell 1997). In fact, in the absence of such trade-offs, polymorphism is not selected for, even under soft selection. Thus despite the experimental attempts cited above, the demonstration - through a proper experimental test - that polymorphism observed under soft selection may be maintained over the long term by selection has yet to be made (Vale 2013).

In principle, demonstrating polymorphism maintenance over the course of an experiment is not sufficient to conclude that environmental heterogeneity selects for polymorphism. It is of prime importance to discern whether observed polymorphisms are due to the negative frequency dependence produced by environmental heterogeneity or not. To do so, one must test for polymorphism protection. Polymorphism is protected if both genotypes increase in frequency when initially rare. Under polymorphism protection, no genotype can ever disappear and polymorphism is expected to be maintained (unless important drift causes the random loss of one genotype).

This article is protected by copyright. All rights reserved.
Here we conducted serial passage experiments reproducing the key stages of soft and hard selection models with bacterial populations (Figure 1C). In order to test the effect of selection regimes on the maintenance of diversity, a local adaptation trade-off was required. Such a pattern was obtained by using two *Escherichia coli* genotypes, one being resistant to tetracycline and the other to nalidixic acid. Populations composed of two genotypes were grown in heterogeneous environments each composed of two different habitats, one containing a very low concentration of tetracycline and the other a very low concentration of nalidixic acid. Low antibiotic concentrations provided a selective advantage to the resistant genotype over the susceptible one, but both genotypes could grow in all conditions. Serial passages were conducted in each heterogeneous environment, as in Bell (1997), to simulate hard and soft selection (Figure 1B). The maintenance of already established polymorphisms (Experiment 1) as well as polymorphism protection (Experiment 2) were assessed over the course of the experiment by using flow cytometry. The duration of the experiment was kept short enough to avoid the emergence of *de novo* mutations (toward a generalist genotype resistant to both antibiotics) and monitoring of selection coefficients *a posteriori* confirmed that no such evolution occurred. Lastly, the diverse effects of selection regimes on polymorphisms were dissected by systematically confronting observed trajectories of genotype frequencies to theoretical predictions. We discuss the potential of such an experimental system to explore the contribution of soft and hard selection to local adaptation polymorphisms.

**Materials and Methods**

**Bacterial strains**

In order to test the effect of selection regimes on the maintenance of polymorphism, a local adaptation trade-off was required. Such a pattern was obtained by using two *E. coli* genotypes, one being resistant to tetracycline and the other to nalidixic acid. The *E. coli* B strains used in this study, REL4548 YFP-Tet® and REL4548 CFP-Nal® derive from the strain REL4548 kindly provided by R. E. Lenski. REL4548 was evolved in Davis minimal (DM) medium supplemented with 25 µg/mL glucose (DM25) for 10,000 generations as part of a long-term evolution experiment (Lenski *et al.* 1991). Gallet *et al.* (2012) then inserted YFP and CFP genes at the rhaA locus of REL4548 using a technique developed by Datsenko and Wanner (2000). A mini-Tn10 derivative 104 (mini-Tnlo tet/Ptac-ATS transposase - Kleckner *et al.* 1991) — was introduced at the *insL-l* locus into REL4548 YFP (clone T121) (Gallet *et al.* 2012) to construct REL4548 YFP-Tet®. The strain REL4548 CFP-Nal® was then created by selecting a resistant REL4548 CFP colony on a LB plate (10 g/L NaCl, 10 g/L tryptone, 5 g/L yeast extract; 15 g agar, 1000 mL H2O) supplemented with 20 µg/mL of nalidixic Acid. These constructs permitted the association of specific antibiotic resistance with a specific fluorescent marker for easy identification. Bacterial strains were stored at -80°C in 15% glycerol stocks.
Serial passages
To conduct the experiments, three heterogeneous environments were created (see Figure 1A and their detailed description in the following section). A heterogeneous environment is composed of two habitats, *i.e.* two independent 50 mL Falcon tubes containing distinct growth media and where bacteria are grown. For each heterogeneous environment, serial passages were conducted following the same protocol. Before the start of the experiment, REL4548 YFP-Tet<sup>8</sup> and REL4548 CFP-Nal<sup>8</sup> genotypes were grown separately overnight in 5 mL of DM25 (37°C, 215 rpm). At T0, the optical density (OD, 600 nm, Eppendorf spectrophotometer) of each culture was measured and a 50-50% mix (in Experiment 1) or a 2-98% mix (in Experiment 2) were made to inoculate the two habitats of each environment (Figure 1B – transfer). After an 18h incubation at 37°C and 215 rpm agitation (Figure 1B – selection), fractions of the bacterial populations from the two habitats composing the environment were mixed (Figure 1B – regulation and pooling). Depending on the selection regime, the regulation step was performed differently: either the same volume (50 µL – hard selection) or different volumes (containing 10<sup>7</sup> cells per habitat – soft selection) from each habitat were pooled in DM0 (*i.e.*, DM medium containing no glucose). During this step, cultures were diluted 10-fold regardless of selection regime. Part of the mix was used to make a glycerol stock (stored at -80°C) for subsequent flow cytometry analysis while the other part was used to inoculate both habitats of the environment of the next passage (50 µL into 5 mL of fresh media – an additional 100-fold dilution).

Heterogeneous environments
Each of the three environments was composed of two habitats, *i.e.* two Falcon tubes containing different growth media: one habitat with a low concentration of tetracycline (Tet) and one habitat with a low concentration of nalidixic acid (Nal). Habitats also differed by their glucose concentration (habitat productivity). As two nutrient concentrations were used (DM2 and DM50 see below), four habitats (*i.e.*, growth media) were used in this study (Nal2, Nal50, Tet2, Tet50). Environment A was composed of habitats Nal2 and Tet50. Environment B corresponded to the pair Nal50 – Tet50. Environment C was composed of habitats Nal50 and Tet2. Since glucose concentrations of the two habitats were equal in Environment B, it was considered as a 'symmetric' environment and both bacterial genotypes were expected to have similar mean fitnesses in the habitat they were adapted to. The two other environments were considered 'asymmetric' as one habitat had a higher glucose concentration than the other expectedly favoring one genotype over the other at the scale of the whole environment.

All media were made on the base of Davis minimal (DM) medium (K_{2}HPO_{4} monohydrate 5.34 g/L, KH_{2}PO_{4} 2 g/L, ammonium sulfate 1 g/L, sodium citrate 0.5 g/L). Bottles were weighed before and after autoclaving and sterile milliQ water was added to compensate for evaporation occurring during sterilization. After autoclaving, media were supplemented with 806 µL/L of
MgSO$_4$•7H$_2$O [1 M], 1 mL/L Thiamine (vitamin B1) [0.2%]. Then, 40 µL/L or 1 mL/L of glucose [2.5%], were added in order to make DM2 and DM50 (2 and 50 µg/mL of glucose being present in the medium, respectively). These media were equivalent to the one used by Lenski et al. (1991), but with different glucose concentrations. Antibiotics were used at subinhibitory concentrations to provide a moderate fitness advantage to the resistant genotype. To take into account week-to-week variations (different medium batch, antibiotic dilution etc.), culture media were tested prior to the start of the experiments, and the relative fitness of bacterial genotypes was measured. Thus, tetracycline and nalidixic acid were added at final concentrations of 0.02 µg/mL and 0.7 µg/mL respectively for the first trial of Experiment 1 and 0.03 µg/mL and 0.8 µg/mL respectively for the second trial of Experiment 1 as well as Experiment 2. These antibiotic concentration adjustments resulted in similar fitnesses in the different experiments (Figure 2A). Finally, to verify that the studied trade-off was not circumvented by the selection of a generalist genotype resistant to both antibiotics, and thus fit to both habitats, we monitored the relative fitness of competitor strains at each transfer. Our results confirm that the relative fitness stayed constant during the entire experiment (Figure S1), and therefore no generalist genotype was selected.

**Experiment 1: Maintenance of established polymorphism**

In Experiment 1, polymorphic populations with initially equal frequencies of both genotypes were grown under hard selection and soft selection regimes. In the first trial, flow cytometer measurements showed that the realized initial frequency of REL4548 YFP-Tet$^R$ was 0.508 and the experiment was conducted over five transfers. The second trial, conducted simultaneously with Experiment 2, started from an initial frequency of REL4548 YFP-Tet$^R$ of 0.437 and had to be interrupted after three transfers due to technical difficulties.

**Experiment 2: Polymorphism protection**

In Experiment 2, the genotype with a global disadvantage in the considered environment (i.e., the genotype with the lowest mean local multiplication rate at the scale of the environment) was initially rare. In Environment A (Nal2-Tet50) the initial frequency of REL4548 YFP-Tet$^R$ was 0.975. In Environment C (Nal50-Tet2) the initial frequency of REL4548 YFP-Tet$^R$ was 0.035. In Environment B (Nal50-Tet50), initially conceived as symmetric, both initial frequencies were tested. The experiment was conducted over two transfer only.

**Flow cytometry**

Flow cytometry allows counting small objects (typically cells) marked with various fluorescent colors. We used a Gallios flow cytometer (Beckman Coulter Inc) - designed to detect small objects such as bacteria - to estimate the relative frequencies of *E. coli* genotypes marked with CFP or YFP. Flow cytometers do not measure the volume of samples they use, precluding the calculation of cell concentrations directly from the counts. Estimating cell concentrations required the addition to the sample of fluorescent beads of a known concentration (AccuCount Fluorescent Particles, 7.0-7.9 µm, Spherotech). Counting fluorescent beads allowed estimation
the samples volume used by the flow cytometer and so estimation of cell concentrations. Results were analyzed with the Kaluza 1.3 software (Beckman coulter Inc). This procedure was performed on overnight cultures and on mixes.

**Statistical analyses**

**Genotypic frequency trajectories**

Observed genotypic frequencies were estimated by counting a minimum of $5 \times 10^4$ cells using flow cytometry. The confidence interval due to measurement error around each frequency was therefore always smaller than 0.004, too small to be represented in Figures 3 and 4. Rather than comparing observed trajectories with time series models, we took advantage of the availability of theoretical predictions (Levene 1953; Dempster 1955) to conduct statistical analyses aimed at testing whether observed trajectories of genotype frequencies over transfers conformed to hard or soft selection models (or none). To do so, we established the predicted trajectories of genotype frequencies for each environment (A, B, C) and each selection regime (hard, soft). Changes in genotype frequencies over transfers under hard and soft selection can be predicted using previously established equations for Levene and Dempster’s models (e.g., available in Appendices 1 and 2 of Ravigné et al. (2004) among others) assuming haploidy, complete and unbiased dispersal, equal habitat frequencies and equal carrying capacities (see the theoretical Appendix of the supporting information section for details). For both models under such conditions, the change in frequency from one transfer to the other only depends on the viability coefficients $W_{i,j}$ of genotype $i$ in habitat $j$. These viability coefficients, also called local fitnesses, correspond to rates of multiplication of each genotype between two transfers. To allow comparing the results with other studies and checking that no generalist genotype was selected over the course of the experiments, the relative fitness (selection coefficients - sensu Chevin 2011) of competitor strains were calculated at each transfer (see Figure S1 in the supporting information section). Selection coefficients ($s$) were calculated as follow:

$$s_{\text{Tet}} = \frac{\ln \frac{C_{\text{Tet}}^{\text{final}}}{C_{\text{Tet}}^{\text{initial}}} - \ln \frac{C_{\text{Nal}}^{\text{final}}}{C_{\text{Nal}}^{\text{initial}}}}{g}$$

where $s_{\text{Tet}}$ is the selection coefficient of the Tet$^R$ genotype, $C_{\text{Tet}}$ and $C_{\text{Nal}}$ correspond respectively to Tet$^R$ and Nal$^R$ genotype concentrations at the beginning (initial) and the end (final) of bacterial growth, and $g$ corresponds to the number of generations.

This article is protected by copyright. All rights reserved.
Viability coefficients for each habitat and each transfer were obtained from flow cytometry measurements as the ratios of cell concentration at the end of the overnight culture to cell concentration at the beginning. Viability coefficients presented non-negligible experimental variability (Figure 2A). To account for this variability, 10,000 trajectories of Tet<sup>R</sup> frequency over transfers were simulated by sampling randomly at each transfer, the viability coefficient values $W_{Tet^R,Nal}$, $W_{Tet^R,Tet}$, $W_{Nal^R,Nal}$, and $W_{Nal^R,Tet}$ among all corresponding values observed over all transfers of Experiment 1. These 10,000 trajectories represented the set of possible trajectories under each environment and each selection regime. The 2.5<sup>th</sup> and 97.5<sup>th</sup> percentile values of the distribution of Tet<sup>R</sup> frequency at each transfer were used to define the 95% envelopes of genotype trajectories under each environment and each selection regime. Similarly, the equilibrium Tet<sup>R</sup> frequency was estimated using the median, 2.5<sup>th</sup> and 97.5<sup>th</sup> percentile values of the distribution of predicted Tet<sup>R</sup> frequencies after 100 transfers. Statistical analysis therefore amounted to checking whether each observed genotype trajectory fell inside or outside the 95% envelopes defined for the corresponding environment.

**Time point comparisons**

Punctual differences between averaged genotypic frequencies between two time points were tested with Generalized linear models (GLM), with “time” as a categorical explanatory factor, and a quasibinomial error type.

**Results**

**Three heterogeneous environments with clear local adaptation trade-off**

Figure 2A shows the local fitnesses (i.e., between transfer multiplication rates) obtained for each bacterial genotype over all replicates in each of the four habitats. From this and the computation of selective coefficients (available in appendix S1), the existence of three different local adaptation trade-offs could be verified (Figure 2B). It was also confirmed that local fitnesses were similar in the two independent trials of Experiment 1 and that multiplication rates – hence bacteria – did not evolve during the experiment (Figure S1). Theoretical predictions showed that Environment A (Nal<sup>2</sup>-Tet50) was so asymmetric that the fixation of the Tet<sup>R</sup> genotype was expected under both hard and soft selection (Figure 3-A1 right-hand side of the x-axis). The expected dynamics of genotype frequency however differed clearly between hard and soft selection (grey areas in Figure 3-A1). In Environments B (Nal50-Tet50) and C (Nal50-Tet2), soft selection was expected to lead to polymorphism maintenance, while hard selection was expected to lead to the fixation of one of the two genotypes (Tet<sup>R</sup> in Environment B and Nal<sup>R</sup> in Environment C, Figures 3-B1 and C1). In Environment B, the dynamics of genotype frequencies over 5 transfers were hardly distinguishable between hard and soft selection (Figure 3-B1).
Effects of selection regimes on the maintenance of polymorphism

Under hard selection, in both environments with asymmetric habitat productivities (Environments A – Nal2-Tet50 and Environment C – Nal50-Tet2), polymorphism was almost completely lost over the experiment (Figures 3-A2 and C2). In environment A, Tet<sup>R</sup> genotype frequency reached an average of 0.991 ± 0.001 after 3 transfers in the two Experiment 1 trials (n=6 replicates) and 0.999 ± 0.0002 after 5 transfers in the long trial (n=3). In environment C, Tet<sup>R</sup> genotype frequency decreased to 0.032 ± 0.027 after 3 transfers in the two Experiment 1 trials (n=6) and 0.002 ± 0.003 after 5 transfers in the long trial (n=3). The trajectories of genotype frequencies fit well with predictions obtained assuming hard selection and fell outside the 95% envelop of theoretical predictions obtained for soft selection (Figures 3-A2 and C2). In the symmetric environment (Environment B), as predicted, polymorphism was almost unchanged at the end of the experiment with only a slight increase of Tet<sup>R</sup> frequency (Figure 3-B2). On average, Tet<sup>R</sup> genotype frequency reached 0.595 ± 0.055 after 3 transfers in the two trials of Experiment 1 (n=6 replicates) and 0.593 ± 0.059 after 5 transfers in the long trial (n=3).

Under soft selection, genetic polymorphism was maintained throughout the experiment regardless of habitat productivities (Figures 3-A3, B3, and C3). In Environment A (Nal2-Tet50), the frequency of the Tet<sup>R</sup> genotype increased at a rate compatible with predictions obtained under soft selection and not with predictions obtained under hard selection (Figure 3-A3). In this environment, although the expected final outcome of selection was the same under hard and soft selection regimes (fixation of the Tet<sup>R</sup> genotype), the rate of evolution was much slower under soft selection than under hard selection.

In Environments B and C where stable polymorphism was expected, genotype frequencies fit well with predictions obtained under soft selection (Figures 3-B3 and C3). In the five-transfer trial, the frequency of the Tet<sup>R</sup> genotype finally attained 0.484 ± 0.045 (expected value: 0.488 with 95% envelope [0.353-0.628]) in Environment B and 0.284 ± 0.033 (expected value: 0.333 with 95% envelope [0.206-0.470]) in Environment C.

Polymorphism protection observed in soft but not in hard selection

The trajectories of genotype frequency observed in Experiment 2 were again in agreement with theoretical expectations (Figure 4). In Environment A (Figure 4A), as observed previously, neither hard selection nor soft selection produced an advantage for the rare genotype. In Environment B the Tet<sup>R</sup> was predicted to go to fixation under hard selection whatever its initial frequency, suggesting a global fitness advantage over the Nal<sup>R</sup> competitor (Figure 3-B1). Under soft selection, Tet<sup>R</sup> frequency nevertheless significantly decreased from a high initial starting value (Figure 4C, GLM with quasibinomial error comparing Tet<sup>R</sup> frequency at T0 and T1: T
value = -3.017, p-value = 0.0393). Similarly in Environment C (Figure 4B), where the TetR genotype had a global fitness disadvantage, TetR frequency significantly increased when initially rare under soft selection only (GLM with quasibinomial error comparing TetR frequency at T1 and T2: T value = 3.458, p-value = 0.0259). In these two environments the observed polymorphism was therefore protected under soft selection.

**Discussion**

From the seminal debate between Levene (1953) and Dempster (1955), theory has suggested that the manner in which populations redistribute among habitats of a given environment is crucial for the long-term maintenance of local adaptation polymorphisms. Under some conditions, soft selection, in which habitat contribution to the next generation is constant, can protect polymorphism by producing negative frequency-dependent selection. In contrast, in the same conditions, hard selection – in which habitat contribution to the next generation varies with habitat genetic composition – do not have such a protective effect on polymorphism and low frequency genotypes are lost (e.g., Christiansen 1974; 1975; Karlin and Campbell 1982; de Meuës et al. 1993). The present experiment is a proof of concept testing for conditions under which the maintenance of a local adaptation polymorphism can be attributed to soft selection.

A prerequisite to study the effect of selection regimes on polymorphism maintenance was to use several well-characterized reproducible trade-offs. In the absence of a local adaptation trade-off, selection is expected to favor a single generalist genotype and environmental heterogeneity cannot lead to stable polymorphism maintenance. The absence of a clear local adaptation trade-off had led previous experimental tests of hard and soft selection to inconclusive results (reviewed in Vale 2013). Here, the use of very low antibiotic concentrations provided a small fitness advantage to resistant genotypes and created a local adaptation trade-off. Since the considered trade-off was not constitutive of bacterial metabolism, it could have been circumvented by the evolution of nalidixic acid resistance in TetR resistant bacteria. Keeping the experiments short (≈50 generations) avoided the selection of a generalist genotype (the monitoring of selection coefficient throughout the experiment confirmed that no generalist genotype evolved in our experiment, see Figure S1), but also precluded the monitoring of polymorphism on the long-term. This latter difficulty, which is actually always a limitation of experimental selection, was overcome by testing for polymorphism protection in a separate experiment (Experiment 2).

In mathematical models of soft and hard selection, population mixing and redistribution among habitats are assumed to happen at each generation and local selection within habitats is assumed to act on viability, i.e., either fecundity or survival rate. Here, transfers were controlled to reproduce the density-regulation steps characteristic of hard and soft selection. But no control was applied during bacterial growth (i.e. between transfers). Thus, for 8 to 10 generations, complex population dynamics or density-dependence could occur within habitats. For instance, priority effects, where the most frequent genotype transforms the environment in
a way that is not favorable to the competitor genotype, could have established and led to frequency dependence within habitats. Furthermore, bacterial populations reached their carrying capacity before transfers, so that competition could be at work not only during the exponential growth phase, but during the stationary phase as well. Congruently, although habitats were all made simultaneously from the same medium preparation, important variability in local multiplication rates was observed. Hence, contrary to original models of hard and soft selection, local multiplication rates (but not fitnesses) varied across transfers and replicates. For instance, the variability of local multiplication rates across transfers made it impossible to use the local multiplication rates observed at the first transfer to predict the dynamics of genotype frequencies over the whole experiment. It also impeded the distinction between trajectories produced under hard and soft selection in an important range of observations (superposed 95% envelopes). Lastly, it translated into unexpected dynamics in Experiment 2 testing for polymorphism protection. Indeed, we observed (i) a transient decrease of the genotype with the highest mean local multiplication rate under hard selection, and symmetrically (ii) an initial decrease of the rare genotype under soft selection (Figure 4C and B). This variability, not present in the models, is actually expected to be even greater in natural conditions. Here we showed that it is possible to account for variable local multiplication rates by combining the models with a bootstrap procedure. This statistical approach revealed that experimental results were in remarkable agreement with theoretical predictions in all treatments. Thus, we can conclude that the impact of hard and soft selection on polymorphism maintenance is robust to a variety of biological processes occurring between regulation steps.

Under the conditions investigated, hard selection never protected polymorphism. The fixation of the genotype with the highest mean local multiplication rate at the scale of the environment was observed within three transfers in the two asymmetric environments. Importantly, in the symmetric environment, polymorphism was still observed after five transfers under hard selection (Figure 3-B2). But deviations to frequencies theoretically expected under soft selection (Figure 3-B2) and the polymorphism protection experiment (Figure 4C) confirmed that such polymorphism consisted of transient polymorphism not being easily removed because of very similar initial frequencies and local multiplication rates. In contrast, under soft selection, polymorphism was never lost over the course of the experiment, even in asymmetric environments where the low fitness genotype was inoculated at a very low relative frequency. The results from Experiment 2 and our theoretical predictions concurred that such polymorphism (1) was only transient in one of the two asymmetric environments (Environment A), and (2) was effectively protected by the existence of a systematic advantage of the rare (i.e., negative frequency dependence) in the two other environments (Environment B and C). Interestingly, in the environment C under soft selection (Figure 3-C3), the genotypic relative frequencies seem to have reached the equilibrium point predicted by theory. Lastly, Experiment 1 confirmed that with all else being equal, even when soft selection is expected to lead to the

This article is protected by copyright. All rights reserved.
fixation of a single genotype (i.e., when its mean fitness at the scale of the environment is very high – environment A in Figure 3), soft selection leads to a slower rate of evolution (sensu modification of genotype frequencies) than hard selection (as shown by Whitlock 2002).

These results illustrate that observing a polymorphism over long periods of time is not sufficient to conclude that it is maintained by soft selection. Under hard selection, transient polymorphism may take long to be removed when habitats are symmetric but it cannot be maintained indefinitely. If one genotype has a slightly higher mean local multiplication rate across habitats than the other, its frequency will increase transfer after transfer up to fixation. The results also confirm that under soft selection, negative frequency-dependent selection emerges even when a genotype has a higher mean fitness at the scale of the environment (Environment C). Indeed, the local regulation step occurring at each transfer opposes the effect of within-habitat selection and hampers invasion of the whole environment by the genotype adapted to the most productive habitat. For very asymmetric trade-offs (and when habitats have different carrying capacities), regulation is not sufficient to generate negative-frequency dependence, which results in polymorphism loss. Overall, the present experiment therefore confirmed the main theoretical predictions on the effects of hard and soft selection on a system that departs from theoretical models by allowing complex population dynamics within habitats between transfers.

From a theoretical perspective it is understood that the conditions for polymorphism maintenance under soft selection are rather stringent (Prout 1968; Christiansen 1974; Maynard Smith and Hoekstra 1980). Various processes, such as drift and mutation, may reduce the range of parameters (trade-off shapes and habitat frequencies) where polymorphism is protected. Polymorphism protection also depends on an equilibrium between local fitnesses (i.e., trade-off shape) and habitat frequencies and carrying capacities. When habitats have contrasted frequencies or carrying capacities even fewer trade-offs shapes allow polymorphism protection. This suggests that soft selection may not be that frequent in nature and that most observed polymorphism is either transient or maintained by other frequency dependence mechanisms (de Meeûs et al. 2000).

In natura, soft selection can be suspected to occur based on species' life cycle. For instance, in some solitary insect parasitoids, a fixed number of individuals, often one, can emerge from each host individual where eggs have been laid (Mackauer 1990). Density regulation is local and the contribution from each infected host to the next generation is fixed. On similar bases, Chao and coll. (2000) argued that soft selection could be the rule rather than the exception in pathogens that kill their host. The rationale is as follows. Pathogens infecting the same host (habitat) constitute a local population. Pathogen genotypes with highest within-host accumulation (highest absolute local fitness) kill the host too quickly to transmit efficiently. The contribution of these genotypes to the next generation is thus not proportional to their absolute fitness but to their fitness relative to other genotypes infecting the same host. Following the same reasoning,
hard selection should be common in pathogens in which virulence is disconnected from transmission rate, for instance for those that either transmit after host death (e.g. Ebola infecting humans or baculoviruses infecting insects), or that do not kill their host, such as the plant pathogenic fungus *Mycobotryum violaceum*, which replaces plant pollen grains by its own spores. Hard selection is also likely frequent in serial passage experiments when parasite transmission is simulated by experimenters. Logically, most serial passage experiments lead to a decrease or disappearance of the initially present polymorphism (for review, see Ebert 1998). Based on our work and previous studies, we advocate that the importance of soft selection in shaping standing genetic variation should not be overlooked (Agrawal 2010; Reznick 2016). In recent experiments, the ‘softness’ of selection (i.e., the contribution of soft selection) was measured in experimental populations of both *Drosophila melanogaster* with different genes (Laffafield et al. 2010; Ho and Agrawal, 2014) and on seedling emergence in *Brassica rapa* (Weis et al. 2015). In addition to highlighting an unexpected sensitivity of softness to genes, individuals and population densities, it was found in both cases that the softness of selection was generally high, cementing the idea that soft selection shapes natural variation at local adaptation loci (Agrawal 2010; Reznick 2016).

Examining the softness of selection is not an easy task as it requires time-consuming experiments implying the manipulation of genotype frequencies, something difficult to apply in the wild. The present work shows that this difficulty can be relieved by modeling the trajectories of genotype frequencies in some parameter ranges and when the variability in local multiplication rates is properly accounted for. Such an approach may help distinguishing between trajectories of genotype frequencies that are compatible with hard selection from those that are compatible with soft selection.

Finally, the present experimental work provides new perspectives for further testing theoretical predictions about the effect of spatial heterogeneity on polymorphism maintenance. The genetics of local adaptation used (two resistance alleles at two different loci) hampers the study of the emergence of local adaptation polymorphisms by gradual evolution. But the experimental system is likely relevant for questions regarding the maintenance of already existing polymorphisms, as observed for species diversity in communities established in heterogeneous environments. For instance, one could test more systematically for the effect of habitat carrying capacities and trade-off intensities on polymorphism maintenance. The present experiment conservatively considered full migration between the two habitats. Migration intensity (Maynard Smith 1966), timing (Ravigné et al. 2004; Débarre and Gandon 2011; Massol 2013), and bias (density dependent migration or habitat selection, de Meûs et al. 1993) affect the range of conditions favorable to polymorphism maintenance and could be tested through a similar experimental design. Estimating the relative importance of spatial and temporal

This article is protected by copyright. All rights reserved.
variability of the environment in shaping polymorphism could also help our understanding of ecological specialization (Massol 2013).

References


Levene, H. 1953. Genetic equilibrium when more than one niche is available. Am. Nat. 87:331–333.


**Figure**

**Figure 1. Experimental design.** Panel A illustrates the composition of the three environments (Env A, B and C) used in our study. Each environment is composed of two habitats (i.e. two independent Falcon tubes), containing either tetracyclin (T-tagged circles) or nalidixic acid (N-tagged circles). Circles diameter represent habitats productivity (small and large circles for DM2 and DM50 respectively). Black and white dots inside of these circles represents the *E. coli* Tet<sup>R</sup> and Nal<sup>R</sup> genotypes respectively. Panel B shows the transfer protocols used under hard (dark grey) and soft selection (light grey) for the asymmetric environment C (Tet2-Nal50). A transfer starts by the inoculation of a bacterial population mix in the two habitats composing the environment. Selection occurs within habitats during the 18 hours of bacterial culture at 37°C. The amount of cells transferred during the regulation step depends on the selection treatment. Under hard selection fixed volumes (50 µL) of each habitat are pooled together in the mixing tube, while under soft selection fixed numbers of cells (10<sup>7</sup> cells) from each habitat are pooled. These mixes are then used to inoculate fresh habitats in the next transfer. Panel C shows the experimental design of both Experiments 1 and 2. Like in panel B, Dark and light grey rectangles represent environments under hard and soft selection, respectively. Like in panel A, the diameter of the circles represents habitat productivity while letters inside the circles symbolize the antibiotic present in the habitat. As Experiment 1 was designed to study the maintenance of a pre-existing polymorphism and Experiment 2 to study polymorphism protection, these experiments were initiated with a 50-50% and a 2-98% mix of bacterial competitors, respectively. In each experiment, experimental treatments were replicated thrice. The entire Experiment 1 was also replicated twice.

This article is protected by copyright. All rights reserved.
Figure 2: Habitats, environments and local adaptation trade-offs. Local fitness is defined as the rate of multiplication between transfers within one habitat (thereby corresponding to viabilities in Levene’s model). Panels: A. Distributions of the local multiplication rates of Nal\(^R\) and Tet\(^R\) genotypes in the long (white whisker charts) and the short (black whisker charts) trials of Experiment 1 for the four habitats used. B. Mean local multiplication rates of Nal\(^R\) (open circles and dashed line) and Tet\(^R\) (filled circles and plain line) genotypes in the three environments used (with confidence intervals over all replicates and transfers of Experiment 1).
Figure 3: Evolution of genotype frequencies under hard and soft selection in experiment 1. **Left-hand panels** show the 95% envelopes of theoretically predicted frequencies the $\text{Tet}^R$ genotype under hard (dark grey) and soft selection (light grey) in environments A (Nal2-Tet50), B (Nal50-Tet50) and C (Nal50-Tet2). Dashed grey lines show the medians of theoretical frequencies. At the right of the x-axis break, theoretically predicted equilibrium frequencies are shown. **Central and right-hand panels** confront these predictions to experimental results obtained under hard and soft selection regimes respectively. Filled circles and lines show the frequencies observed in the three transfers of the short trial. Open circles and dashed lines correspond to the five transfers of the long trial. Note that initial frequencies slightly differ between the two trials.
Figure 4: Polymorphism protection. All panels show the dynamics of the frequency of the Tet<sup>R</sup> genotype over successive transfers in Experiment 2. Frequencies observed under hard and soft selection are represented with filled circles connected by solid lines and opened circles connected by dashed lines, respectively. Dark and light grey areas show the 95% envelopes of theoretically predicted frequencies under hard and soft selection, respectively. Panels: A. Environment A (Nal2-Tet50) with initial high Tet<sup>R</sup> genotype frequency. B. Environment C (Nal50-Tet2) with initial low Tet<sup>R</sup> genotype frequency. C. Environment B (Nal50-Tet50) with initial high Tet<sup>R</sup> bacteria genotype frequency. D. Environment B with initial low Tet<sup>R</sup> genotype frequency. In each panel, the inset shows a magnification of the frequency trajectories.