Peer Community In Evolutionary Biology

Experimental test of the conditions of maintenance of polymorphism under hard and soft selection

Stephanie Bedhomme *based on peer reviews by* ***Joachim Hermisson*** *and 2 anonymous reviewers*

Romain Gallet, Rémy Froissart, Virginie Ravigné (2017) Things softly attained are long retained: Dissecting the Impacts of Selection Regimes on Polymorphism Maintenance in Experimental Spatially Heterogeneous Environments. Missing preprint_server, ver. 3, peer-reviewed and recommended by Peer Community in Evolutionary Biology.
https://doi.org/10.1101/100743

Submitted: 17 January 2017, Recommended: 03 April 2017

Cite this recommendation as:
Bedhomme, S. (2017) Experimental test of the conditions of maintenance of polymorphism under hard and soft selection. *Peer Community in Evolutionary Biology*, 100020. 10.24072/pci.evolbiol.100020

Published: 03 April 2017

Theoretical work, initiated by Levene (1953) [1] and Dempster (1955) [2], suggests that within a given environment, the way populations are regulated and contribute to the next generation is a key factor for the maintenance of local adaptation polymorphism. In this theoretical context, hard selection describes the situation where the genetic composition of each population affects its contribution to the next generation whereas soft selection describes the case where the contribution of each population is fixed, whatever its genetic composition. Soft selection is able to maintain polymorphism, whereas hard selection invariably leads to the fixation of one of the alleles. Although the specific conditions (e.g. of migration between populations or drift level) in which this prediction holds have been studied in details by theoreticians, experimental tests have mainly failed, usually leading to the conclusion that the allele frequency dynamics was driven by other mechanisms in the experimental systems and conditions used. Gallet, Froissart and Ravigné [3] have set up a bacterial experimental system which allowed them to convincingly demonstrate that soft selection generates the conditions for polymorphism maintenance when hard selection does not, everything else being equal. The key ingredients of their experimental system are (1) the possibility to accurately produce hard and soft selection regimes when daily transferring the populations and (2) the ability to establish artificial well-characterized reproducible trade-offs. To do so, they used two genotypes resisting each one to one antibiotic and combined, across habitats, low antibiotic doses and difference in medium productivity. The experimental approach contains two complementary parts: the first one is looking at changes in the frequencies of two genotypes,

initially introduced at around 50% each, over a small number of generations (ca 40) in different environments and selection regimes (soft/hard) and the second one is convincingly showing polymorphism protection by establishing that in soft selection regimes, the lowest fitness genotype is not eliminated even when introduced at low frequency. In this manuscript, a key point is the dialog between theoretical and experimental approaches. The experiments have been thought and designed to be as close as possible to the situations analysed in theoretical work. For example, the experimental polymorphism protection test (experiment 2) closely matches the equivalent analysis classically performed in theoretical approaches. This close fit between theory and experiment is clearly a strength of this study. This said, the experimental system allowing them to realise this close match also has some limitations. For example, changes in allele frequencies could only be monitored over a quite low number of generations because a longer time-scale would have allowed the contribution of de novo mutations and the likely emergence of a generalist genotype resisting to both antibiotics used to generate the local adaptation trade-offs. These limitations, as well as the actual significance of the experimental tests, are discussed in deep details in the manuscript.

References:

[1] Levene H. 1953. Genetic equilibrium when more than one niche is available. American Naturalist 87: 331–333. doi: [10.1086/281792](https://doi.org/10.1086/281792)

[2] Dempster ER. 1955. Maintenance of genetic heterogeneity. Cold Spring Harbor Symposia on Quantitative Biology. 20: 25–32. doi: [10.1101/SQB.1955.020.01.005](https://doi.org/10.1101/SQB.1955.020.01.005)

[3] Gallet R, Froissart R, Ravigné V. 2017. Things softly attained are long retained: dissecting the impacts of selection regimes on polymorphism maintenance in experimental spatially heterogeneous environments. bioRxiv 100743; doi: [10.1101/100743](https://doi.org/10.1101/100743)

Reviews

Evaluation round #2

DOI or URL of the preprint: **10.1101/100743**
Version of the preprint: 2

Authors' reply, 31 March 2017

We now have carefully taken all additional corrections into account. Please find the modified version. We sincerely wish to thank all three reviewers for their valuable contribution to our paper.

Decision by Stephanie Bedhomme, posted 01 April 2017

Decision of 2nd round

Two of the reviewers have now read the revised version of your manuscript and were generally happy with the changes made. One of them, however, suggested minor revisions that have to be incorporated in the manuscript. Assuming that these changes will be done, I will recommend the manuscript.

Reviewed by Joachim Hermisson, 01 April 2017

I had raised two points in my previous report. The first and main point concerned a more precise formulation of claims and results. The authors have done this in their revision. The second point concerned the discussion of future use of the experimental setup. The authors have added a comment on that. I had also mentioned the short duration of the experiment (50 generations), a point that has been made much stronger by the other referees. Here, I agree with the authors that the action of soft/hard selection (with its potential - or not - to maintain polymorphism) has been demonstrated. The invadability experiments make a convincing case (for me) for an internal equilibrium because of selection against rare types. Lack of long-term stability has only been an issue (for me) as it limits the use of the setup for further experiments. It seems that this is indeed an issue that cannot be easily overcome. However, the authors clearly mention this and so (for me) this point has also been taken care of. I have no further concerns.

Reviewed by anonymous reviewer 2, 01 April 2017

I am overall satisfied with this new version of the manuscript. I only have a few minor comments - Last paragraph of the introduction, first sentence: I would remove “under some conditions”. Either you explain what you mean precisely by this statement or it is useless. - Last paragraph of the introduction, last sentence: replace “for polymorphisms to maintain” by “for polymorphisms to be maintained”. - Page 5, first paragraph, first column: replace “cytometer measurement” by “cytometer measurements”. - Page 5, first paragraph, second column: I would replace “even the genotype with a lower absolute fitness cannot be eliminated from the population hampering the polymorphism to erode” with “the genotype with the lowest absolute fitness will be eliminated from the population, thereby fostering the maintenance of polymorphism”. - Page 9, first paragraph: “Disposing of” in English means “eliminating”... I guess you mean something else here... Also, this was not exactly a ‘pre-requisite’, but a finding in its own right, I would say... I would rephrase to “identifying a (...) was an important finding of this study”. - Page 9, first paragraph, 2nd column: replace “was bound to maintain” by “will probably be maintained”

Evaluation round #1

DOI or URL of the preprint: **10.1101/100743**
Version of the preprint: 1

Authors’ reply, 20 March 2017

Reply to reviewers’ comments

Thank you for submitting your work to PCI Evol Biol. Three reviewers have now read your manuscript and all agree that the question is highly interesting and the experimental work of great quality. However all have concerns for diverse reasons, partially linked, on what is actually demonstrated by the results. I would thus ask you to revise your manuscript and particularly the discussion to address the different concerns pointed by the reviewers. Two of them have also suggested minor changes or corrections that have to be included. I look forward to seeing your revised manuscript.

RF, RG and VR: I would like to thank all three reviewers for very relevant comments that we hope, helped improving the manuscript.

Reviewed by anonymous reviewer, 2017-02-10 17:01

R1: In the manuscript « Maintenance of Polymorphism in Spatially Heterogeneous Environments: An Experimental Test of Hard and Soft Selection”, the authors aim at testing experimentally an old theoretical prediction using a bacterial system. The introduction provides a clear overview of the knowledge on the central question of the paper and the previous attempts to test this prediction, the experimental design and the results

are clearly presented and the paper is generally well written. However, even though the experimental design accurately reproduce the two selection regimes (hard/soft) for initially mixed populations composed of two genotypes having a selective advantage in different habitats of the environment, I do not think that the author can claim that they have shown "maintenance of polymorphism" because the experiment was conducted on five transfers (40 generations) at most. What they have shown is that the ingredients for the maintenance of polymorphism are present and experiment 2 on frequency-dependence is convincing to me on this point.

A longer-term experiment would be necessary to really show "maintenance" and it is on the longer run that deviation from the theoretical predictions might appear. It is actually very likely that performing more transfers with this biological material would have conducted to the de novo appearance and selection of nalidixic acid resistance in the TetR genotype. Indeed, resistance to nalidixic acid is mainly due to "change of target" point mutations in the gyrase gene. This would lead to a double resistant, generalist across environments. This suggests that the biological system chosen is not ideal because it requires to restrain the experiment in time. Additionally, the fact that the two genotypes chosen to start the experiment differ at two loci (resistance to Tet and resistance to Nal) and are not a wild-type and a mutant (differing only at one locus) is changing the potential contribution of de novo mutations: whereas in a wild-type/mut system, mutation at the focus loci are only changing the frequency of the two genotypes, in a system such as the one chosen by the authors, de novo mutation at one of the two focus loci is generating a third genotype, potentially fitter than the two initial ones. More generally, if the goal is to test the theoretical prediction taking into account only the frequency changes of the two genotypes in a context where new mutations do not arise, bacteria might not be the ideal system as the time scale to convincingly show polymorphism maintenance is a time scale at which new mutations will occur in large bacterial population: not only double resistant but also compensatory mutations reducing the cost of resistance, changing the fitness of each initial genotype in each habitat.

RF, RG and VR: Questions regarding the experimental system and our conclusions are important points raised by R1 and the other reviewers. As reported in the manuscript, previous attempts to demonstrate the effect of soft selection on the maintenance of polymorphism have been unsuccessful, mainly because of the difficulty to work with a clear stable local adaptation trade-off. Finding the right biological system and the right conditions to test this theory is tricky. Reviewer 1 has actually produced a very good summary of the constraints imposed by the experimental system we set up. We however disagree on a number of crucial points.

First and foremost, we actually think that *under the trade-offs studied* long-term maintenance of polymorphism was proven in both environments B and C. Demonstrating that an equilibrium genotype frequency has been reached and is indefinitely maintained is actually difficult. Observing a polymorphism over a "large" number of generations is insufficient to prove that polymorphism will further maintain. As noted by R1, there is no guarantee that "something" different occurs after the end of the experiment. The way indefinite polymorphism maintenance is proven in theoretical models is by testing for polymorphism protection. One computes the change in frequency when each genotype is initially rare. If both changes are positive, then both genotypes increase in frequency when initially rare, implying that they cannot disappear, and that the polymorphism is bound to maintain. Polymorphism protection is also called "advantage of the rare" or "negative frequency dependence". This is exactly what Experiment 2 has shown. In Environment B (resp. Environment C), the genotype NalR (resp. TetR) was less fit at the scale of the environment. As illustrated by the hard selection treatment, it was bound to disappear in the absence of a mechanism countering the effect of selection. But at low frequency NalR (resp. TetR) is advantaged in soft selection and increases in frequency, meaning that it cannot be eliminated from the population. Therefore, as long as the trade-off is unaltered, soft selection promotes polymorphism maintenance.

Second, we agree with the statement of R1 that deviation from the prediction may appear in case of effective additional transfers. But as noted by R1, such deviation (if any) would certainly be provoked by factors (such as mutation) that would deviate from the model we aimed at testing, that is the hard-

soft selection regimes. We fully agree that disposing of an experimental system where the trade-off would naturally stay unaltered would be more elegant/powerful, but up to now, despite numerous (published and unpublished) attempts, no one has succeeded in that task. There is a trade-off (again!) between the necessity to study a large number of generations and the rate of evolution of studied species. Bacteria offer the possibility to study tens of generations within a reasonable amount of time (the same experiment would last 3 years on Drosophila) and not to have to deal with drift. The paper is clearly not about bacterial evolution, nor about the robustness of local adaptation patterns in front of evolution. The aim was really to test whether all else being equal, a pulse of local density regulation once every few generations is sufficient to favor polymorphism maintenance. Third, we did not employ polymorphism maintenance in a strict long-term context. At the scale of the experiment, soft selection already promoted polymorphism maintenance. In Environment A, it was shown that the rate of fixation was slower under soft selection than under hard selection. Thus within the timeframe of the experiment, more (transient) polymorphism was observed. In Environment C, fixation was observed under hard selection and not under soft selection. And in that case the polymorphism observed, not significantly different from its final expected value, was shown to be supported by negative frequency dependence. This is sufficient to claim that soft selection promotes polymorphism maintenance over tens of generations by two different mechanisms: slower rate of evolution and/or negative frequency dependence.

We now have tried to (i) clarify the prominent role of Experiment 2 in proving indefinite polymorphism maintenance under stable trade-offs, (ii) explicitly discuss the limitations of the system to study evolutionary mechanisms other than polymorphism (iii) systematically distinguish between "short-term" (i.e., within the duration of the experiment) and "long-term" polymorphism maintenance (i.e., beyond the experiment). This is visible in the end of introduction (page 3), in the description of experiment 2 (page 6) and in the discussion (page 9).

R1: Minor points: Some experimental choice (or constraints) should be justified or discussed more clearly: - How were the antibiotic concentrations chosen?

RF, RG and VR: Antibiotic concentrations were chosen empirically. Antibiotic concentrations were tuned in order to provide a moderate selective advantage to the resistant genotype over the susceptible one. Too much antibiotic and the susceptible genotype does not grow, too low and the resistant genotype has no selective advantage.

R1: Why are they different for experiment 1 and 2? The difference is rather small such that I do not think that it prevents from establishing links between the results of the two experiments but the fact that there is a difference is intriguing to me.

RF, RG and VR: Good question. There are some inevitable week-to-week variations due to the preparation of a new batch of culture medium, different antibiotic dilutions or unknown factors. To make sure that those variations would not impact our experiments, we tested all media before using them, measured the relative fitness of the strains under these conditions, and adjusted the antibiotic concentrations in order to have the same relative fitnesses among different replicates. We now explicitly explain this rationale in the material and methods (Mat & Meth, end of page 3).

R1: Why do the two trials of experiment 1 have a different number of transfers?

RF, RG and VR: Simply because we could not conduct the second trial any longer due to technical issues. We decided to keep it because it is an independent replicate of the experiment (it was done 2 years after the first trial) and because over the 3 valid transfers the genotype frequency changes are extremely congruent with those observed in the first trial, suggesting a high reproducibility of the experiment. We now give a brief mention of these technical issues in the material and methods (Mat & Meth, end of page 5).

Reviewed by Joachim HERMISSON, 2017-02-10 17:01

R2: Theoretical models already from the 1950s have demonstrated that spatially heterogeneous selection

can be a potent force to maintain genetic variation in a population (Levene 1953, Dempster 1955). These models also show that genetic variation is only preserved in the face of gene flow if density regulation in the structured population is local (so-called soft selection) rather than global (hard selection). There is of course ample evidence of genetic variation in spatially structured populations under heterogeneous selection in nature. However, it is often difficult to demonstrate the mode of density regulation (hard or soft selection) and to rule out alternative explanations for the maintenance of variation. To bridge this gap between a theoretical model and empirical observation, Gallet et al have set up an evolution experiment with E. Coli in a heterogeneous environment, exposed to either hard and soft selection. Their observations from the experiment are in perfect agreement with the predictions from the theoretical model.

The accomplishment of the ms is that the authors have shown that they have constructed an experimental setup where that can clearly demonstrate hard and soft selection because all other confounding (and potentially relevant) factors are under control. In addition, all relevant parameters can be measured in a way that we get more or less full quantitative agreement of the experimental results with the simple theoretical model. Some people call this a "smelly computer". However, experimental systems are often a pain and it is not easy to create a setup that is under sufficient control that we fully understand its dynamics and even get quantitative agreement with a model. The promise of such a system is that we can gradually add further natural factors to test the importance of each of them for the dynamics and the potential of the system to maintain genetic variation.

RF, RG and VR: We fully agree with Dr Hermisson's comment. In particular we recognize the "smelly computer" side of the experiment and look forward to being able to further use the system to test the relative importance of different factors in maintaining polymorphism.

R2: This said, the system might be a bit limited since the number of generations tested is not large. It is larger generation numbers where problems usually begin.

RF, RG and VR: True. Please refer to our answers to R1, who shared the same concerns. This point is now fully addressed in the discussion section that was significantly rewritten.

R2: My main concern with the ms is the following: What the authors claim to do is to test a theoretical prediction. Namely, that soft selection in a heterogeneous environment will maintain polymorphism (under some additional conditions that are precisely described in a model), while hard selection won't. The authors do find this in their experimental system and conclude that "this is the first conclusive experimental demonstration ... ". However, I do not think that this is what their experiment is able to do. Imagine, for a moment, that their results would have been in stark contrast with the predictions of the theoretical model. What would have been the conclusion? That - other than previously thought - soft selection does not maintain genetic variation (or hard selection does)? Well, certainly not. The conclusion would have been that relevant factors other than hard or soft selection dominate the dynamics observed in the experiment. Maybe the results do give us an indication that soft / hard selection is indeed a factor that is potent and relevant enough that it is not dominated by other factors that cannot be controlled – at least for E. Coli experimental evolution. This is nice to know, but it is not the result that the ms claims to have shown. In other words: What we can test is not the conceptual insight of the model ("soft selection preserves variation, hard selection does not"), but only a model prediction for a specific empirical system ("the dynamics of this particular system can be fully explained by soft/hard selection"). In a revision, the authors should formulate their claims much more precisely, restricting them to the results that their experiment is able to show.

RF, RG and VR: We globally agree with this comment. We changed the title, deleted the corresponding sentences and clarified our claim: in the experiment, despite uncontrolled aspects of bacterial population dynamics not present in Levene (1953) and Dempster (1955) models, a pulse of soft selection once every few generations was sufficient to trigger polymorphism maintenance. This is visible is the end of the introduction (page 3) and in the discussion (page 11).

R2: If they want to present their system as a tool for future experimental work, they should add a paragraph detailing its promises.

RF, RG and VR: The last paragraph of the discussion section (page 11-12) lists possible uses of this system to explore how various factors may interact in shaping local adaptation polymorphisms, e.g., temporal variability vs. spatial heterogeneity.

Reviewed by anonymous reviewer, 2017-02-10 17:01

R3: The manuscript aims at testing the classical prediction that polymorphisms are better maintained in populations submitted to soft selection than in populations under hard selection. To this aim, they have done a few serial passages of polymorphic bacteria populations, in which individual bacteria were either resistant to tetracycline or to nalidixic acid, in environments composed of two patches, each with one of these antibiotics, either at high or low concentrations. Passages were done according to hard (fixed volume, variable number) or soft selection (fixed number, variable volume) on bacteria. Then, they have contrasted the frequency of bacteria with each genotype along the passages with a simulation based upon the theoretical models underlying hard and soft selection. They found that bacteria dynamics are compatible with such theoretical predictions. In addition, they also showed that predictions on how these selection regimes affect frequency-dependent selection are also met. Overall, I found the article well written and presenting interesting findings (although N=3 is a bit low...). Still, I would like to have seen more discussion on the potential long-term consequences of these findings. Indeed, the authors claim that they have used a small number of generations to ensure that no generalist would arise through novel mutations. However, isn't whether such findings will be robust in the face of other strategies precisely the question? I think that this should at least be discussed.

RF, RG and VR: For a full answer to this comment, please read our answer to R1 who brought up this point as well. In brief, we claim that we evidenced that indefinite polymorphism maintenance was guaranteed assuming that the trade-off stays unaltered (Experiment2). Although very interesting, whether the trade-off stays unaltered is really not the question here (actually there is a high probability that a generalist genotype evolves in the longer run). The question was whether given a trade-off, soft selection would promote polymorphism maintenance as predicted by theory. This said, we now explicitly discuss this limitation of the experimental system and the questions that it is designed to treat.

R3: Also, I personally would not call the pattern observed a 'trade-off'. It is certainly a pattern of local adaptation, but there is no evidence of antagonistic pleiotropy. For sure, each researcher may have his/her own definition of a trade-off. However, it must be clear that different definitions imply different predictions. Indeed, in this 'loose' definition, it is relatively easy to generate a clone that has both mutations, conferring resistance to the two antibiotics (i.e., a generalist). In contrast, if there would be antagonistic pleiotropy between those two mutations, the emergence of a generalist would be less likely. Again, it would be interesting to extend the discussion on this topic.

RF, RG and VR: Good point. We agree with R3 that the trade-off used is artificial. It is neither the result of antagonistic pleiotropy nor due to the accumulation of deleterious mutations. As noted by R3 (and R1 and R2), it is thus bound to be altered by the occurrence of de novo mutations. We however do not authorize evolution by novo mutations in the experiment. The trade-off is a fixed constrained, imposed by experimenters instead of bacterial metabolism, as a "negative genetic correlations in fitness across different habitats, Levins 1962". We discuss this point in the discussion. But we prefer not to elaborate on how evolution by de novo mutation would affect this trade-off because, as noted by all three referees, then the experimental system would not be the best option and other trade-offs would be more interesting to study (but see beginning of discussion page 10)

R3: I also have a few minor comments:

- Explain "and thus habitats themselves contribute to the next generation".
  RF, RG and VR: Done

- Explain "due to decreases in host accumulation".
  RF, RG and VR: Done

- Replace “considering selection regime could be important for analyses” by “considering that these selection regime could be important”.
  RF, RG and VR: Done

- Explain “see Bell and Reboud 1997 in which unexpected selection was suspected to have played a role”.
  RF, RG and VR: Done

- Figure 1A should be referred to before figure 1b.
  RF, RG and VR: Done

- Figure 1: consider having all panels in A and B with the same scale….
  RF, RG and VR: We now have all panels in B with the same scale. We however prefer to keep panels A as they were because differences between habitats would be almost invisible in the upper panels if they were on a 0-5000 scale.

- Equation (3): is it really pt/2? It does not seem logical to me…
  RF, RG and VR: Yes, it is. That’s because after mixing half the populations goes in habitat 1 and the other half in habitat 2 and that both habitats have the same carrying capacities. In the generic model, the term is ci/(c1+c2) where ci is the carrying capacity of habitat i, i.e., the number of individuals that survive local density regulation in habitat i.

Decision by Stephanie Bedhomme, posted 16 February 2017

Revision needed

Thank you for submitting your work to PCI Evol Biol. Three reviewers have now read your manuscript and all agree that the question is highly interesting and the experimental work of great quality. However all have concerns for diverse reasons, partially linked, on what is actually demonstrated by the results. I would thus ask you to revise your manuscript and particularly the discussion to address the different concerns pointed by the reviewers. Two of them have also suggested minor changes or corrections that have to be included. I look forward to seeing your revised manuscript.

Reviewed by anonymous reviewer 1, 10 February 2017

In the manuscript « Maintenance of Polymorphism in Spatially Heterogeneous Environments: An Experimental Test of Hard and Soft Selection”, the authors aim at testing experimentally an old theoretical prediction using a bacterial system. The introduction provides a clear overview of the knowledge on the central question of the paper and the previous attempts to test this prediction, the experimental design and the results are clearly presented and the paper is generally well written.

However, even though the experimental design accurately reproduce the two selection regimes (hard/soft) for initially mixed populations composed of two genotypes having a selective advantage in different habitats of the environment, I do not think that the author can claim that they have shown “maintenance of polymorphism” because the experiment was conducted on five transfers (40 generations) at most. What they have shown is that the ingredients for the maintenance of polymorphism are present and experiment 2 on frequency-dependence is convincing to me on this point.

A longer-term experiment would be necessary to really show “maintenance” and it is on the longer run that deviation from the theoretical predictions might appear. It is actually very likely that performing more transfers with this biological material would have conducted to the de novo appearance and selection of nalidixic acid resistance in the TetR genotype. Indeed, resistance to nalidixic acid is mainly due to “change of target” point mutations in the gyrase gene. This would lead to a double resistant, generalist across environments. This suggests that the biological system chosen is not ideal because it requires to restrain the experiment in time. Additionally, the fact that the two genotypes chosen to start the experiment differ at two loci (resistance to

Tet and resistance to Nal) and are not a wild-type and a mutant (differing only at one locus) is changing the potential contribution of de novo mutations: whereas in a wild-type/mut system, mutation at the focus loci are only changing the frequency of the two genotypes, in a system such as the one chosen by the authors, de novo mutation at one of the two focus loci is generating a third genotype, potentially fitter than the two initial ones. More generally, if the goal is to test the theoretical prediction taking into account only the frequency changes of the two genotypes in a context where new mutations do not arise, bacteria might not be the ideal system as the time scale to convincingly show polymorphism maintenance is a time scale at which new mutations will occur in large bacterial population: not only double resistant but also compensatory mutations reducing the cost of resistance, changing the fitness of each initial genotype in each habitat.

Minor points: Some experimental choice (or constraints) should be justified or discussed more clearly: - How were the antibiotic concentrations chosen? - Why are they different for experiment 1 and 2? The difference is rather small such that I do not think that it prevents from establishing links between the results of the two experiments but the fact that there is a difference is intriguing to me. - Why do the two trials of experiment 1 have a different number of transfers?

Reviewed by Joachim Hermisson, 10 February 2017

Theoretical models already from the 1950s have demonstrated that spatially heterogeneous selection can be a potent force to maintain genetic variation in a population (Levene 1953, Dempster 1955). These models also show that genetic variation is only preserved in the face of gene flow if density regulation in the structured population is local (so-called soft selection) rather than global (hard selection). There is of course ample evidence of genetic variation in spatially structured populations under heterogeneous selection in nature. However, it is often difficult to demonstrate the mode of density regulation (hard or soft selection) and to rule out alternative explanations for the maintenance of variation. To bridge this gap between a theoretical model and empirical observation, Gallet et al have set up an evolution experiment with E. Coli in a heterogeneous environment, exposed to either hard and soft selection. Their observations from the experiment are in perfect agreement with the predictions from the theoretical model.

The accomplishment of the ms is that the authors have shown that they have constructed an experimental setup where that can clearly demonstrate hard and soft selection because all other confounding (and potentially relevant) factors are under control. In addition, all relevant parameters can be measured in a way that we get more or less full quantitative agreement of the experimental results with the simple theoretical model. Some people call this a ”smelly computer”. However, experimental systems are often a pain and it is not easy to create a setup that is under sufficient control that we fully understand its dynamics and even get quantitative agreement with a model. The promise of such a system is that we can gradually add further natural factors to test the importance of each of them for the dynamics and the potential of the system to maintain genetic variation. This said, the system might be a bit limited since the number of generations tested is not large. It is larger generation numbers where problems usually begin.

My main concern with the ms is the following: What the authors claim to do is to test a theoretical prediction. Namely, that soft selection in a heterogeneous environment will maintain polymorphism (under some additional conditions that are precisely described in a model), while hard selection won’t. The authors do find this in their experimental system and conclude that ”this is the first conclusive experimental demonstration ... ”. However, I do not think that this is what their experiment is able to do. Imagine, for a moment, that their results would have been in stark contrast with the predictions of the theoretical model. What would have been the conclusion? That - other than previously thought - soft selection does not maintain genetic variation (or hard selection does)? Well, certainly not. The conclusion would have been that relevant factors other than hard or soft selection dominate the dynamics observed in the experiment.

Maybe the results do give us an indication that soft / hard selection is indeed a factor that is potent and relevant enough that it is not dominated by other factors that cannot be controlled – at least for E. Coli experimental evolution. This is nice to know, but it is not the result that the ms claims to have shown. In

other words: What we can test is not the conceptual insight of the model ("soft selection preserves variation, hard selection does not"), but only a model prediction for a specific empirical system ("the dynamics of *this particular system* can be fully explained by soft/hard selection"). In a revision, the authors should formulate their claims much more precisely, restricting them to the results that their experiment is able to show. If they want to present their system as a tool for future experimental work, they should add a paragraph detailing its promises.

Reviewed by anonymous reviewer 2, 10 February 2017

The manuscript aims at testing the classical prediction that polymorphisms are better maintained in populations submitted to soft selection than in populations under hard selection. To this aim, they have done a few serial passages of polymorphic bacteria populations, in which individual bacteria were either resistant to tetracycline or to nalidixic acid, in environments composed of two patches, each with one of these antibiotics, either at high or low concentrations. Passages were done according to hard (fixed volume, variable number) or soft selection (fixed number, variable volume) on bacteria. Then, they have contrasted the frequency of bacteria with each genotype along the passages with a simulation based upon the theoretical models underlying hard and soft selection. They found that bacteria dynamics are compatible with such theoretical predictions. In addition, they also showed that predictions on how these selection regimes affect frequency-dependent selection are also met. Overall, I found the article well written and presenting interesting findings (although N=3 is a bit low...). Still, I would like to have seen more discussion on the potential long-term consequences of these findings. Indeed, the authors claim that they have used a small number of generations to ensure that no generalist would arise through novel mutations. However, isn't whether such findings will be robust in the face of other strategies precisely the question? I think that this should at least be discussed. Also, I personally would not call the pattern observed a 'trade-off'. It is certainly a pattern of local adaptation, but there is no evidence of antagonistic pleiotropy. For sure, each researcher may have his/her own definition of a trade-off. However, it must be clear that different definitions imply different predictions. Indeed, in this 'loose' definition, it is relatively easy to generate a clone that has both mutations, conferring resistance to the two antibiotics (i.e., a generalist). In contrast, if there would be antagonistic pleiotropy between those two mutations, the emergence of a generalist would be less likely. Again, it would be interesting to extend the discussion on this topic. I also have a few minor comments: - Explain "and thus habitats themselves contribute to the next generation". - Explain "due to decreases in host accumulation". - Replace "considering selection regime could be important for analyses" by "considering that these selection regime could be important". - Explain "see Bell and Reboud 1997 in which unexpected selection was suspected to have played a role". - Figure 1A should be referred to before figure 1b. - Figure 1: consider having all panels in A and B with the same scale.... - Equation (3): is it really pt/2? It does not seem logical to me...