Peer Community In
Evolutionary Biology

# Unmasking the delusive appearance of negative frequency-dependent selection

***Ignacio Bravo*** *based on peer reviews by* ***David Baltrus*** *and 2 anonymous reviewers*

Dustin Brisson (2017) Negative frequency-dependent selection is frequently confounding. Missing preprint_server, ver. Missing article_version, peer-reviewed and recommended by Peer Community in Evolutionary Biology. 10.1101/113324

Submitted: 03 March 2017, Recommended: 07 July 2017

**Cite this recommendation as:**
Bravo, I. (2017) Unmasking the delusive appearance of negative frequency-dependent selection. *Peer Community in Evolutionary Biology*, 100024. 10.24072/pci.evolbiol.100024

Published: 07 July 2017

---

Explaining the processes that maintain polymorphisms in a population has been a fundamental line of research in evolutionary biology. One of the main mechanisms identified that preserves genetic diversity is negative frequency-dependent selection (NFDS), which constitutes a powerful framework for interpreting the presence of persistent polymorphisms. Nevertheless, a number of patterns that are often explained by invoking NFDS may also be compatible with, and possibly more easily explained by, different processes. In the present manuscript [1], Brisson acknowledges first that genuine NFDS has been instrumental for our understanding on the dynamics that perpetuate polymorphisms, and that the power and importance of NFDS cannot be disregarded. Second, the author aims at identifying certain of the processes that may result in maintenance of genetic diversity, and whose outcome may be mistaken for NFDS, namely directional selection in changing environments, density-dependent fitness, multiple niche selection and community diversity. The author claims that systematic resort to NFDS as explanatory device may have lead to its application to systems where it does not apply or that do not fulfil the basic assumptions of NFDS. The author struggles in the text to provide with a precise, verbal definition of NFDS, and the exchanges with the reviewers during the recommendation process show that agreeing on such a verbal definition of NFDS is not trivial. Probably a profound mathematical formulation of the varying value of a genotype's fitness relative to other competing ones as a function of their frequency (developing further the synthesis by Heino [2]) may still be wanting. Indeed, the text is intended for a broad audience of evolutionary biologists with operational mathematical knowledge and interest in models, rather than for modellers or biomathematicians. Nevertheless, the manuscript is rich in references to original literature, elaborates on interesting lines of thought and discussion and will hopefully trigger novel experimental and formal research to clarify the role of NFDS and to discern between alternative mechanisms that may render similar patterns of maintenance of genetic diversity.

***References:***

[1] Brisson D. 2017. Negative frequency-dependent selection is frequently confounding. bioRxiv 113324, ver. 3 of 20th June 2017. doi: [10.1101/113324](https://doi.org/10.1101/113324)

[2] Heino M, Metz JAJ and Kaitala V. 1998. The enigma of frequency-dependent selection. Trends in Ecology & Evolution 13: 367-370. doi: [1016/S0169-5347(98)01380-9](https://doi.org/10.1016/S0169-5347(98)01380-9)

# Reviews

## Evaluation round #2

DOI or URL of the preprint: **10.1101/113324**
Version of the preprint: 2

### Authors' reply, 20 June 2017

Recommendation for preprint "Negative frequency-dependent selection is frequently confounding". by Ignacio Bravo, 2017-06-17 22:43 Manuscript: https://doi.org/10.1101/113324 Decision & reviews Three experts in the field have provided with extensive, high-quality and fair scientific judgement on the revised version of th text. All three reviewers agree on the pertinence and timeliness of the subject, and all of them also agree on the need of sound review highlighting the false friends that may be mistaken by negative frequency-dependent selection. Nevertheless, two of the reviewers consider that some of the key terms used in the text are ill-defined, and that precisely because the aim of the review is to clarify and draw clear borders, semantics (i.e. meanings) are in this case of the utmost importance. I largely agree with this perception. Further, I understand that the author aims at a target audience that does not expect a hard mathematical description. Nevertheless, the experts' opinion remains that a sounder mathematical foundation will undoubtedly contribute to a richer and deeper manuscript. Finally, I think that the suggestion to provide clear published examples of such misinterpretations of NFDS may also be illuminating, but I understand that expurgating the literature may not be in the interest nor in the aim of the author. Globally, I consider that the manuscript is potentially very interesting, and I would suggest to revise it considering the comments provided by the authors, before it can be recommended by PCI Evol Biol.

–Thank you and the reviewers for taking the time to review this manuscript. I have changed the manuscript in accordance with the reviewers' suggestions by, most importantly, clarifying the key terms identified by the reviewers. A consideration of the target audience with regard to this manuscript was noted by multiple reviewers and the editor. In this regard, I agree with the editor and the reviewers who noted that including formal descriptions of the models has the potential to alienate the very audience that could be most impacted by these types of concept articles. It is important to note that this paper is neither building nor critiquing the mathematical models that have been previously formalized, an important topic that is beyond the goal of this paper. The primary issue being addressed is the heuristic application of the concept of negative frequency dependent selection to polymorphisms in biological systems in which the general concept does it apply, as noted by the reviewers (most explicitly reviewer 2). The types of errors noted here are not being committed by math-bio researchers who, despite being the primary reviewers of this manuscript, are thus not the target audience. Critical evaluations of the mathematical models underlying negative frequency dependent selection has been done elsewhere (although more work is certainly necessary). Given the target audience, it is my opinion that re-describing the math underlying these models, which can be found in the cited work, will be counter-productive as (1) the models themselves are not being critiqued here, (2) no new models are being proposed, and (3) the primary issue is that the general concepts in those models is being misunderstood by those that do not dive into the math-bio weeds. Lastly, while I agree that providing examples of errors in the literature can be illuminating, it is likely to be detrimental to scientific progress in general by turning off the very audience that could benefit from this paper (I also think it is mean).

Reviewed by anonymous reviewer, 2017-06-17 22:43

The author has made some marginal changes to its original manuscript, but my main source of disagreement remains. However, I do not consider the paper is wrong, as our disagreement mostly revolves around semantics, so I do not have further requests for changes. I will however provide a brief reply to the author's response to my comments, as well as some minor formal suggestions on the text.

The main issue I have with the paper is that negative frequency-dependence is defined very narrowly: according to the author, "rare variants must be advantageous because of their relatively rarity, regardless of the ecological interaction that mediates it". With this definition, negative frequency-dependence will clearly be a very special case, and it is no surprise that the author can find many exceptions. At the other end of the spectrum, frequency dependence has been defined more broadly, with reference to the dimension of the environmental feedback (Heino et al 1998, Metz & Geritz 2016). With this broader definition, the distinctions pointed out by the author are not so relevant. So, whether or not one finds the author's contribution useful depends on where one sets the cursor.

–I believe the reviewer and I agree: negative frequency dependent selection is a narrow selective force that has nevertheless been applied broadly in a heuristic manner and often incorrectly. The point of the paper is that the negative frequency dependent selection framework, in the narrow sense, is regularly described as the force maintaining a natural polymorphism when it clearly is not. I agree with the reviewer that many of the issues dealt with here are relevant in different ways (or not at all) when discussing the dependence of natural selection on frequency of alleles in a broad sense, but these are outside the scope of this paper.

Personally, I don't think a narrow definition of frequency-dependence is useful. Unless we look at very simple models, I don't see how fitness can depend on alele frequency only, and not on population densities or other environmental variables. This is the main reason why I was (and remain) rather unconvinced by the author's essay. However, I may be wrong, and since there is scope for debate, I think the author's contribution should play a role in this debate.

–It is important to note that I am not attempting to define frequency dependence narrowly, but to describe a specific slice of frequency dependence that, in some areas of evolutionary biology, is important and is often invoked. An additional detail - negative frequency dependent selection does not require that relative fitness is a function of allele frequency only, but that allele frequency must be part of the function (that is, there must be a strict dependence on frequency). I have tried to clarify this in the text. I agree with the reviewer that this manuscript, which discusses the areas to which this concept should apply, needs to be openly discussed and debated.

A few additional remarks: • I don't agree with the interpretation of the paper by Kisdi 1999. In fact, by definition, when the population sits at a fitness minimum, any rare variant can invade, resulting in evolutionary branching. For me this qualifies as negative frequency-dependence, although it does not fit, if I understand correctly, the author's definition.

–I agree with the reviewer about the definition of a fitness minimum. But the conditions of the model presented in the paper prevent invasion due to free recombination (according to the interpretation of the authors). This is a topic that could be discussed at length, although beer will likely be necessary for any real progress.

• About short-term and long-term evolution: my remark was that, for long-term evolution, you also need to account for the genetic variation due to new mutations, not only for the standing genetic variation.

–I see. I agree with the reviewer about the importance of this topic but think this is beyond the general scope for target audience.

• The author uses the concepts of hard and soft selection to discuss density- and frequency-dependent selection. At the risk of passing for an iconoclast, I do not think these concepts, which have their origin in very specific models that are best viewed as caricatures, are really empirically or theoretically very useful. But this is best left for another discussion!

–I would very much like to not discuss hard or soft selection in this manuscript.

Some minor comments on the text: • top of p. 4: ”the overwhelming majority OF this broad field”?

–done

• bottom of p. 6: it could be useful to explain in detail how this is different from the self-incompatibility allele example.

–I am not sure how to make this more clear, but welcome suggestions. The main concept I was trying to make clear is that this example does not rely on relative rarity.

• I think there are too many ”luminaries” in this text. Some rewording could be welcome lest the reader be dazzled!

–done

Reviewed by David Baltrus, 2017-06-17 22:43

I have read and reviewed the revised version of this manuscript, and have also read and taken into account the well reasoned critiques of other reviewers.

To me, the revised manuscript does read as more precise in its language and more limited in scope (almost to a fault, but not quite). I can definitely see the anonymous reviewer’s points and agree that these are quite relevant critiques. This is certainly a situation that can be semantically challenging, and I think it’s also going to be the case where specific viewpoints, backgrounds, and research foci are going to influence how the article is viewed.

To me, this article is well placed and hits the right notes for an audience that has heard the words ”frequency dependent selection” but which are not necessarily truly experts in that topic. It’s an article that will do well, especially given the examples, to clear up some confusion across research topics that might seek to invoke the phrase ”frequency dependent selection”. I get the feeling that this is the intent, and I think the article does deliver in this context and it will make people think. Although I thought the first version was pretty easy to digest, this version does set up the contrasts better and lays out the necessary details more clearly than the first.

In the context of my own research background, and with the acknowledgement that others might not see it this way, I think this article will be well received by the audiences at the intended levels. I would encourage others to read it at a broad level, but with the point of view that some of the specific details could be debatable. I don’t view this as a problem, but as an admonition that nature and evolutionary scenarios can be complicated. This is actually an improvement from many of the discussions/invocations of frequency dependent selection I’ve seen lately.

Reviewed by anonymous reviewer, 2017-06-17 22:43

(First, I have to declare that I did not have access to the first version of this manuscript, therefore I am not commenting on the previous revision round.)

The understanding of a non trivial phenomenon such as biological evolution requires identifying the processes that underlie it. A sharp, unambiguous circumscription of these processes is however needed so that research remains able to reveal yet undetected causes of observed effects, thus allowing theory to be improved, not to mention the accuracy of its prediction.

In the present work, the author discusses the alleged key role of negative frequency-dependent selection (NFDS) in persistent genetic variation and how a maintained polymorphism could be explained by other processes – namely directional selection with environmental change, density-dependent selection, multiple niche selection and community diversity – that could be misinterpreted as NFDS.

Both the topic and the purpose of this paper are of significant interest in the field of evolutionary biology and I can only support its ambition to disentangle NFDS from concurrent selective contexts that can as well lead to stable genetic diversity – one of the main questions in evolutionary biology. I find the text interesting, richly documented and I really think the reflection it brings can be helpful to a large community of theoretical and field or experimental evolutionary biologists, mainly because it tries to link conceptual thinking to concrete examples, and as such, I fully agree with the conclusion paragraphs.

Nonetheless, in my opinion, the present version of the manuscript fails to reach its praiseworthy goal. The reason why I cannot suggest its recommendation as it stands is mainly because the central principle of NFDS is still loosely defined and consequently prevents further rigorous discussion. This is in particular due to the lack of formalism that could have greatly contributed to the needed clarification and lead to compelling proof. As exposed by [1], mathematical models act as "proof-of-concept" tests of verbal explanations, the scope of which is otherwise restricted to trivial cases.

I do not claim that this paper should turn into a heavy mathematical demonstration (it may even be self-defeating as part of that format could repel part of the aimed audience). I nevertheless suggest the author to be very neat on the definitions, to state the assumptions and to try to translate them into simple mathematical conditions in a way similar to e.g. reference 36 of the paper (Heino et al. 1998), while clearly indicating the conceptual novelty (with respect to NFDS) the manuscript brings compared to this same reference (which already contains, technically, key arguments that should be helpful for the present discussion – see Table 1 of Heino et al. 1998).

–I agree with the reviewer that Heino et al did a very good job describing frequency dependence to the target audience of mathematically-oriented evolutionary biologists and the mathematical conditions were well presented. I also agree with the reviewer that the addition of a more mathematical demonstration will repel the very audience that this manuscript should reach. In the revised manuscript, I have pointed to the underlying models where possible for those that would like to discuss the mathematical details of the models. I would also point out that this paper does not describe any conceptual novelty with respect to negative frequency dependent selection. The descriptions in this article draw directly from the assumptions and underlying foundations of published models.

It is also important to remember that the primary objective of this paper is not to describe the key role of negative frequency-dependent selection in persistent genetic variation per se but to point out that negative frequency dependent selection is often used as a heuristic to describe genetic variation in systems where it does not apply as they violate the key assumption/condition of the process. The issue at hand is not a critique of prior modeling nor is it that mathematically-oriented research has made errors that need to be rectified. Further, a critical evaluation of the mathematical models underlying negative frequency dependent selection has been done well and, while more rigorous discussion of model details is always necessary, this manuscript is not intended spark those discussions.

Precisely, when the author introduces NFDS in its paper, it is not clear if he does acknowledge this definition as the sentence starts with "In models..." (and further "In these models..."). But more importantly, it is not clear if the fact that "rare variants have a selective advantage specifically because etc." is part of the definition of NFDS, one of its necessary conditions, or a sufficient condition. The same ambiguity holds for the next sentence "Thus, NFDS can maintain genetic polymorphisms...", while I wonder if "can" implies potentiality or systematic ability.

–I have made these definitions more clear throughout.

First, I am unsatisfied by the lack of definition of "Darwinian fitness" (p.3) (how is it computed, is it absolute, relative?) and their related notions of "per capita fitness" (p.8), "per capita growth rate" (p.11), "per capita rate of increase" (p.11) "per capita reproductive advantage" (p.5), "per capita fitness advantage" (p.6), "per capita selective advantage" (p.6), the author inconsistently uses throughout the paper.

–I agree with the reviewer that the multitude of terms can be confusing. I have limited the number of terms and defined them throughout the manuscript.

Second, if I assume that what the author means by fitness is equivalent to the instantaneous growth rate (as done in classical frameworks that link evolution to population dynamics), that is to say that one can write dn/dt = w.n, where n is the density of the focal genotype and w if the so-called fitness, therefore it is not sufficient that w is a decreasing function of the frequency of the focal genotype p = n/N (N being the total population density) for the selection to be negative frequency-dependent. Indeed, if we take e.g. w1 = 2 – p and w2 = 1 – q, with q = 1 – p, then both fitnesses are decreasing function of the respective frequency of each genotype but the genotype 1 will always increase in frequency even where frequent, as the following short R script can show:

```
library(”deSolve”) ODE1=function(t,y,parms=NULL)\list(c((2-y[1]/(sum(y)))y[1],(1-y[2]/(sum(y))y[2])))\ Y=lsoda(c(1,10),Times,ODE1
; plot(Y[,1],Y[,2]/(Y[,2]+Y[,3]),type='l',lwd=2,ylim=c(0,1)) Y=lsoda(c(10,1),Times,ODE1) ; plot(Y[,1],Y[,2]/(Y[,2]+Y[,3]),type='l',lwd=2,ylim=
```

As a result, the fitness decrease with frequency is not a sufficient condition of NFDS (I acknowledge that an ecological mechanism that would generate the fitness functions as assumed here seems unrealistic but this then highlights the fact that fitness cannot be any function of frequency and therefore it must be stated in the definition).

–I agree with the reviewer that it is not sufficient that that w is a decreasing function of the frequency and I apologize for the confusion. Negative frequency dependent selection occurs when relative fitness (not absolute fitness) is a function of relative abundance. While this is a necessary condition for negative frequency dependent selection, it is not sufficient for a stable polymorphism. However, it is beyond the scope of this paper to review the conditions necessary to achieve a stable polymorphism through negative frequency dependent selection as (1) the issue at hand is that natural polymorphisms are identified and then explained (incorrectly) using negative frequency dependent selection in the literature and (2) this has been done elsewhere.

Third, the fitness decrease with frequency is neither a necessary condition of NFDS. Indeed, one can imagine w1 to be a constant but w2 to decrease with frequency, e.g. w2 = 1/(2q). The following short R script then shows that despite the fact that the fitness of the focal genotype (1) is frequency-independent, its frequency increases if rare because of it rarity and decreases when frequent because of its commonness and that polymorphism is maintained, although the causality here is mechanistically indirect:

```
library(”deSolve”) ODE2=function(t,y,parms=NULL)\list(c(y[1],0.5/(y[2]/sum(y))*y[2]))\ Y=lsoda(c(1,10),Times,ODE1)
; plot(Y[,1],Y[,2]/(Y[,2]+Y[,3]),type='l',lwd=2,ylim=c(0,1)) Y=lsoda(c(10,1),Times,ODE1) ; plot(Y[,1],Y[,2]/(Y[,2]+Y[,3]),type='l',lwd=2,ylim=
```

–Sorry again for the confusion, the explanation to this is similar that given above. Relative fitness decreases with increasing frequency is a necessary condition of negative frequency dependent selection, absolute fitness is not. I have made this more clear in the manuscript.

As a conclusion, the definition of NFDS provided in this paper is logically flawed. Rather, one should investigate the following mathematical condition derived from the replicator's equation and little calculus (making plain the "advantageous"/"disadvantageous" dichotomy by the way): NFDS <=> (w > w' if p~0) and (w < w' if p~1) where w is the fitness of the focal genotype the frequency of which is p and w' is the fitness of the other genotype(s) (if several, an appropriate weighted mean must be taken). If one does agree with this formulation, therefore attesting or contesting NFDS in a given eco-evolutionary context should be reduced to the estimation of four fitness limits.

–I agree with the reviewer that additional work can be done on the modeling of negative frequency dependent selection, which is well beyond the scope of this paper. I encourage the reviewer to create a formal rubric that one could apply to data or models to determine the applicability of negative frequency dependent selection in a given system. This would be an important addition to the literature and could help clear up these issues in a very practical way. Here, I have made the definition more clear to avoid the misunderstanding between relative and absolute fitness as it is described in these models.

Apart from the NFDS definition, I think the author should insist on the literature that has mistaken the concurrent processes chosen by the author for NFDS instead of focusing on few great names of evolutionary biology. More precisely, it is well established in the vast biomathematical literature that Lotka-Volterra systems are not examples of NFDS. In addition, from the quote provided, it is quite unclear that Lewontin (not "Lewinton", p.8) was thinking about NFDS rather than density-dependence.

–I agree with the reviewer that pointing out mistakes in the literature can be illuminating, but it is often counter-productive because it can alienate the very audience one hopes to reach (I also consider it to be mean). I agree that Lewontin was thinking about density-dependence. However, this section of the book is explicitly about negative frequency dependent selection (the preceding paragraphs include “Rare-genotype advantage is an attractive hypothesis” and “a rare allele will increase in frequency but will not become fixed in the population, because as it gets commoner the fitness of its carriers decreases, and the other alleles are now favored.” That he was thinking of density dependence while discussion negative frequency dependent selection is indeed the point of this section.

My next comments are minor and hereafter given linearly: p.6 – “A prominent … lines of thought”. This historical digression sounds a bit off topic.

–I have found this a useful literary device and would prefer to keep it.

p.7 - “The presence or frequency …”. It is worth to mention that this does not hold for all parasites and requires cross-reactivity, within-host competition and epidemiological feedback to be negligible.

–this example is specifically about flu and is used only to illustrate of the ideas presented. Adding the complications of other systems could add confusion and obscure the primary point.

p.8 – “the logic suggests”. The implied logic is not straightforward for me.

–I agree, the logic is incorrect, but is what was stated in the literature. Describing the incomplete logic is the point of this section.

p.8 – “is always at 100%”. Better “equal to 1”?

–For the target audience, I believe 100% is more clear

p.9 – A system brace and a centered label would be more aesthetic for Eq.1.

–ok

p.9 – Low frequency and low density are more explicit than “relatively rarity” and “numerical rarity”.

–For the target audience and to agree with the other language to avoid confusion, I think these terms are more clear.

Figure 1. Perhaps mention, for the sake of completeness, that the self-incompatibility investigated here is gametophytic. “Pollinated” instead of “pollenated”? The legend seems written for two alleles while three alleles are shown in the graphs.

–I have fixed the legend such that the third allele does not feel left out

Figure 2. The graphs should be labeled. The legends of panels A and C do not indicate what the quantity corresponding to the y-axis is. “at 100% frequencies” is unclear.

–I have added more descriptions to the figures to clarify

Figure 3. The graphs should be labeled and indicated in the legend. The first sentence of the legend should be rephrased. The parameter values can be given in the supplementary material only.

Supplementary material. The outline of this appendix is inconsistent. I have checked the R scripts and they work fine, producing the Figures (1 and 3) provided in the main text. A word is missing at the end of “by percent each is found in the”. Use “seedling” instead of “baby plant”/”babies”? Only the first plot command is relevant for Figure 1. As for Figure 3, the script lacks some comments within the for loops, and the final plot. –Agreed.

Summary: I request a substantial revision of this manuscript, essentially with respect to the definition and formalization of the NFDS and an emphasis on the novelty the discussion here provided compared to previous literature.

Reference [1] Servedio MR, Brandvain Y, Dhole S, Fitzpatrick CL, Goldberg EE, et al. (2014) Not Just a Theory—The Utility of Mathematical Models in Evolutionary Biology. PLOS Biology 12(12): e1002017. https://doi.org/10.1371/journal.pbio.1002017

## Decision by Ignacio Bravo, posted 21 June 2017

**Recommendation for preprint ”Negative frequency-dependent selection is frequently confounding”.**

Three experts in the field have provided with extensive, high-quality and fair scientific judgement on the revised version of th text. All three reviewers agree on the pertinence and timeliness of the subject, and all of them also agree on the need of sound review highlighting the false friends that may be mistaken by negative frequency-dependent selection. Nevertheless, two of the reviewers consider that some of the key terms used in the text are ill-defined, and that precisely because the aim of the review is to clarify and draw clear borders, semantics (i.e. meanings) are in this case of the utmost importance. I largely agree with this perception. Further, I understand that the author aims at a target audience that does not expect a hard mathematical description. Nevertheless, the experts' opinion remains that a sounder mathematical foundation will undoubtedly contribute to a richer and deeper manuscript. Finally, I think that the suggestion to provide clear published examples of such misinterpretations of NFDS may also be illuminating, but I understand that expurgating the literature may not be in the interest nor in the aim of the author. Globally, I consider that the manuscript is potentially very interesting, and I would suggest to revise it considering the comments provided by the authors, before it can be recommended by PCI Evol Biol.

## Reviewed by anonymous reviewer 1, 20 June 2017

The author has made some marginal changes to its original manuscript, but my main source of disagreement remains. However, I do not consider the paper is wrong, as our disagreement mostly revolves around semantics, so I do not have further requests for changes. I will however provide a brief reply to the author's response to my comments, as well as some minor formal suggestions on the text.

The main issue I have with the paper is that negative frequency-dependence is defined very narrowly: according to the author, "rare variants must be advantageous because of their relatively rarity, regardless of the ecological interaction that mediates it". With this definition, negative frequency-dependence will clearly be a very special case, and it is no surprise that the author can find many exceptions. At the other end of the spectrum, frequency dependence has been defined more broadly, with reference to the dimension of the environmental feedback (Heino et al 1998, Metz & Geritz 2016). With this broader definition, the distinctions pointed out by the author are not so relevant. So, whether or not one finds the author's contribution useful depends on where one sets the cursor.

Personally, I don't think a narrow definition of frequency-dependence is useful. Unless we look at very simple models, I don't see how fitness can depend on alele frequency only, and not on population densities or other environmental variables. This is the main reason why I was (and remain) rather unconvinced by the author's essay. However, I may be wrong, and since there is scope for debate, I think the author's contribution should play a role in this debate.

A few additional remarks:

- I don't agree with the interpretation of the paper by Kisdi 1999. In fact, by definition, when the population sits at a fitness minimum, any rare variant can invade, resulting in evolutionary branching. For me this qualifies as negative frequency-dependence, although it does not fit, if I understand correctly, the author's definition.

- About short-term and long-term evolution: my remark was that, for long-term evolution, you also need to account for the genetic variation due to new mutations, not only for the standing genetic variation.

- The author uses the concepts of hard and soft selection to discuss density- and frequency-dependent selection. At the risk of passing for an iconoclast, I do not think these concepts, which have their origin in very specific models that are best viewed as caricatures, are really empirically or theoretically very useful. But this is best left for another discussion!

Some minor comments on the text:

- top of p. 4: "the overwhelming majority OF this broad field"?

- bottom of p. 6: it could be useful to explain in detail how this is different from the self-incompatibility allele example.

- I think there are too many "luminaries" in this text. Some rewording could be welcome lest the reader be dazzled!

## Reviewed by David Baltrus, 20 June 2017

I have read and reviewed the revised version of this manuscript, and have also read and taken into account the well reasoned critiques of other reviewers.

To me, the revised manuscript does read as more precise in its language and more limited in scope (almost to a fault, but not quite). I can definitely see the anonymous reviewer's points and agree that these are quite relevant critiques. This is certainly a situation that can be semantically challenging, and I think it's also going to be the case where specific viewpoints, backgrounds, and research foci are going to influence how the article is viewed.

To me, this article is well placed and hits the right notes for an audience that has heard the words ”frequency dependent selection” but which are not necessarily truly experts in that topic. It's an article that will do well, especially given the examples, to clear up some confusion across research topics that might seek to invoke the phrase ”frequency dependent selection”. I get the feeling that this is the intent, and I think the article does deliver in this context and it will make people think. Although I thought the first version was pretty easy to digest, this version does set up the contrasts better and lays out the necessary details more clearly than the first.

In the context of my own research background, and with the acknowledgement that others might not see it this way, I think this article will be well received by the audiences at the intended levels. I would encourage others to read it at a broad level, but with the point of view that some of the specific details could be debatable. I don't view this as a problem, but as an admonition that nature and evolutionary scenarios can be complicated. This is actually an improvement from many of the discussions/invocations of frequency dependent selection I've seen lately.

## Reviewed by anonymous reviewer 2, 20 June 2017

(First, I have to declare that I did not have access to the first version of this manuscript, therefore I am not commenting on the previous revision round.)

The understanding of a non trivial phenomenon such as biological evolution requires identifying the processes that underlie it. A sharp, unambiguous circumscription of these processes is however needed so that research remains able to reveal yet undetected causes of observed effects, thus allowing theory to be improved, not to mention the accuracy of its prediction.

In the present work, the author discusses the alleged key role of negative frequency-dependent selection (NFDS) in persistent genetic variation and how a maintained polymorphism could be explained by other processes – namely directional selection with environmental change, density-dependent selection, multiple niche selection and community diversity – that could be misinterpreted as NFDS.

Both the topic and the purpose of this paper are of significant interest in the field of evolutionary biology and I can only support its ambition to disentangle NFDS from concurrent selective contexts that can as well lead to stable genetic diversity – one of the main questions in evolutionary biology. I find the text interesting, richly documented and I really think the reflection it brings can be helpful to a large community of theoretical and field or experimental evolutionary biologists, mainly because it tries to link conceptual thinking to concrete examples, and as such, I fully agree with the conclusion paragraphs.

Nonetheless, in my opinion, the present version of the manuscript fails to reach its praiseworthy goal. The reason why I cannot suggest its recommendation as it stands is mainly because the central principle of NFDS is still loosely defined and consequently prevents further rigorous discussion. This is in particular due to the lack of formalism that could have greatly contributed to the needed clarification and lead to compelling proof. As exposed by [1], mathematical models act as “proof-of-concept” tests of verbal explanations, the scope of which is otherwise restricted to trivial cases.

I do not claim that this paper should turn into a heavy mathematical demonstration (it may even be self-defeating as part of that format could repel part of the aimed audience). I nevertheless suggest the author to be very neat on the definitions, to state the assumptions and to try to translate them into simple mathematical conditions in a way similar to e.g. reference 36 of the paper (Heino et al. 1998), while clearly indicating the conceptual novelty (with respect to NFDS) the manuscript brings compared to this same reference (which already contains, technically, key arguments that should be helpful for the present discussion – see Table 1 of Heino et al. 1998).

Precisely, when the author introduces NFDS in its paper, it is not clear if he does acknowledge this definition as the sentence starts with “In models...” (and further “In these models...”). But more importantly, it is not clear if the fact that “rare variants have a selective advantage specifically because etc.” is part of the definition of NFDS, one of its necessary conditions, or a sufficient condition. The same ambiguity holds for the next sentence “Thus, NFDS can maintain genetic polymorphisms...”, while I wonder if “can” implies potentiality or systematic ability.

First, I am unsatisfied by the lack of definition of “Darwinian fitness” (p.3) (how is it computed, is it absolute, relative?) and their related notions of “per capita fitness” (p.8), “per capita growth rate” (p.11), “per capita rate of increase” (p.11) “per capita reproductive advantage” (p.5), “per capita fitness advantage” (p.6), “per capita selective advantage” (p.6), the author inconsistently uses throughout the paper. Second, if I assume that what the author means by fitness is equivalent to the instantaneous growth rate (as done in classical frameworks that link evolution to population dynamics), that is to say that one can write
dn/dt = w.n,
where n is the density of the focal genotype and w if the so-called fitness, therefore it is not sufficient that w is a decreasing function of the frequency of the focal genotype p = n/N (N being the total population density) for the selection to be negative frequency-dependent. Indeed, if we take e.g. w1 = 2 – p and w2 = 1 – q, with q = 1 – p, then both fitnesses are decreasing function of the respective frequency of each genotype but the genotype 1 will always increase in frequency even where frequent, as the following short R script can show:

```
    library(”deSolve”)
ODE1=function(t,y,parms=NULL)\list(c((2-y[1]/(sum(y)))y[1],(1-y[2]/(sum(y))y[2])))\
Y=lsoda(c(1,10),Times,ODE1) ; plot(Y[,1],Y[,2]/(Y[,2]+Y[,3]),type='l',lwd=2,ylim=c(0,1))
Y=lsoda(c(10,1),Times,ODE1) ; plot(Y[,1],Y[,2]/(Y[,2]+Y[,3]),type='l',lwd=2,ylim=c(0,1))
```

As a result, the fitness decrease with frequency is not a sufficient condition of NFDS (I acknowledge that an ecological mechanism that would generate the fitness functions as assumed here seems unrealistic but this then highlights the fact that fitness cannot be any function of frequency and therefore it must be stated in the definition).

Third, the fitness decrease with frequency is neither a necessary condition of NFDS. Indeed, one can imagine w1 to be a constant but w2 to decrease with frequency, e.g. w2 = 1/(2q). The following short R script then shows that despite the fact that the fitness of the focal genotype (1) is frequency-independent, its frequency increases if rare *because* of it rarity and decreases when frequent *because* of its commonness and that polymorphism is maintained, although the causality here is mechanistically indirect:

```
    library(”deSolve”)
ODE2=function(t,y,parms=NULL)\list(c(y[1],0.5/(y[2]/sum(y))*y[2]))\
Y=lsoda(c(1,10),Times,ODE1) ; plot(Y[,1],Y[,2]/(Y[,2]+Y[,3]),type='l',lwd=2,ylim=c(0,1))
Y=lsoda(c(10,1),Times,ODE1) ; plot(Y[,1],Y[,2]/(Y[,2]+Y[,3]),type='l',lwd=2,ylim=c(0,1))
```

As a conclusion, the definition of NFDS provided in this paper is logically flawed. Rather, one should investigate the following mathematical condition derived from the replicator’s equation and little calculus (making plain the “advantageous”/”disadvantageous” dichotomy by the way):
NFDS <=> (w > w’ if p~0) and (w < w’ if p~1)
where w is the fitness of the focal genotype the frequency of which is p and w’ is the fitness of the other genotype(s) (if several, an appropriate weighted mean must be taken). If one does agree with this formulation, therefore attesting or contesting NFDS in a given eco-evolutionary context should be reduced to the estimation of four fitness limits.

Apart from the NFDS definition, I think the author should insist on the literature that has mistaken the concurrent processes chosen by the author for NFDS instead of focusing on few great names of evolutionary biology. More precisely, it is well established in the vast biomathematical literature that Lotka-Volterra systems are not examples of NFDS. In addition, from the quote provided, it is quite unclear that Lewontin (not "Lewinton", p.8) was thinking about NFDS rather than density-dependence.

My next comments are minor and hereafter given linearly:
p.6 – "A prominent ... lines of thought". This historical digression sounds a bit off topic.
p.7 - "The presence or frequency ...". It is worth to mention that this does not hold for all parasites and requires cross-reactivity, within-host competition and epidemiological feedback to be negligible.
p.8 – "the logic suggests". The implied logic is not straightforward for me.
p.8 – "is always at 100%". Better "equal to 1"?
p.9 – A system brace and a centered label would be more aesthetic for Eq.1.
p.9 – Low frequency and low density are more explicit than "relatively rarity" and "numerical rarity".
Figure 1. Perhaps mention, for the sake of completeness, that the self-incompatibility investigated here is gametophytic. "Pollinated" instead of "pollenated"? The legend seems written for two alleles while three alleles are shown in the graphs.
Figure 2. The graphs should be labeled. The legends of panels A and C do not indicate what the quantity corresponding to the y-axis is. "at 100% frequencies" is unclear.
Figure 3. The graphs should be labeled and indicated in the legend. The first sentence of the legend should be rephrased. The parameter values can be given in the supplementary material only.
Supplementary material. The outline of this appendix is inconsistent. I have checked the R scripts and they work fine, producing the Figures (1 and 3) provided in the main text. A word is missing at the end of "by percent each is found in the". Use "seedling" instead of "baby plant"/"babies"? Only the first plot command is relevant for Figure 1. As for Figure 3, the script lacks some comments within the for loops, and the final plot.
Summary: I request a substantial revision of this manuscript, essentially with respect to the definition and formalization of the NFDS and an emphasis on the novelty the discussion here provided compared to previous literature.

Reference

[1] Servedio MR, Brandvain Y, Dhole S, Fitzpatrick CL, Goldberg EE, et al. (2014) Not Just a Theory—The Utility of Mathematical Models in Evolutionary Biology. PLOS Biology 12(12): e1002017. https://doi.org/10.1371/journal.pbio.1002017

## Evaluation round #1

DOI or URL of the preprint: **10.1101/113324**
Version of the preprint: 1

### Authors' reply, 17 May 2017

**Decision & reviews**

**Two experts in the field have reviewed the preprint entitled "Negative frequency-dependent selection is frequently confounding". Although the reviewers' impressions on the text are largely divergent, I consider that they remain compatible, and I subscribe (to different extents) to their opinions. Thus, I profoundly agree with the author in the need of a systematic identification of patterns that may be explained by an underlaying process of negative frequence-dependent selection (NFDS), but that may also be compatible and possibly more easily explained by different process. I also acknowledge the effort in referring to the original literature (although in the case of referring to books, I would appreciate if the quote could be identified with more precision, including pages or at least chapters). The author**

**elaborates on four processes that may lead to maintenance of diversity and whose results may be mistaken for NFDS: directional selection in changing environments, density-dependent fitness, multi-niche selection and community diversity. Reviewer#2 has raised a number of important points in the direction of how the width of the definition used for NFDS has an impact on the interpretation/usefulness of the concept itself. While the author has chosen to stick to a classical population genetics definition of NFDS, the reviewer claims that this framework may not be appropriate for analysing the ecological and long-term implications of a genuine NFDS. These conflicting views need to be reconciled. The reviewer criticises then the difference used by the author between relative and absolute abundances, with respect to the carrying capacity of the environment, and its implication. I think that this is a major criticism that also requires to be addressed. The same applies to the appropriateness of the verbal arguments, the completeness of the example used in Figure 3 and the lack of a thorough mathematical analysis. Finally, I also agree with both reviewers that the fourth process "community diversity" is the less-well developed and somehow disconnected with the discussion of allele/variant frequencies in the three previous processes. Globally, I consider that the aim of the review is extremely pertinent and that it may be very useful to clarify when NFDS is truly the most likely explanatory process for a given observed pattern. Nevertheless, a number of important points need to be addressed, and the text and largely benefit from a more explicit introduction of ecology into the population genetics definition of NFDS and from the mathematical and not only verbose systematisation of different patterns being compatible with different underlying processes.**

I appreciate the effort of Drs. Bravo and Baltrus, as well as an anonymous reviewer, for their careful attention and helpful criticism of this preprint that have resulted in a more clear and more accurate manuscript. The main concerns raised focus on (1) the usefulness of analyzing ecologically complex populations through negative frequency dependent selection framework (and potentially population genetic frameworks in general); (2) clarity concerning the distinction between relative abundance and absolute abundance; and (3) greater mathematical descriptions of these processes. The anonymous reviewer correctly notes that the ideas in this paper focus rather narrowly on a small subset of the impacts of frequency on natural selection without explicitly stating this. I have clarified that the objective of this work is to systematically explore patterns that may be explained by an underlying process of negative frequency-dependent selection and those that are compatible and possibly more easily explained by other process. This is done at the expense of a complete treatment of selection as a function of frequency and the associated discussions about hard and soft selection. I contend that the many examples of natural polymorphisms maintained by negative frequency dependent selection that are not controversial demonstrate that this framework is appropriate for analyzing its ecological and long-term implications. I have also clarified through this explicit focus that for negative frequency dependent selection to operate, rare variants must be advantageous because of their relatively rarity, regardless of the ecological interaction that mediates it. Lastly, I have included an analysis of some of the previously published models from which some of the illustrative examples are derived. **Reviewed by anonymous reviewer, 2017-05-09 17:31**

I appreciate the perspective this reviewer has highlighted in response to this manuscript, which has resulted in a much more clear and specific revised manuscript. The primary concerns of the reviewer, as I understand them, are (1) the incomplete treatment of the vast and confusing literature on the impact of allele frequency and population density on the strength of natural selection and (2) a limited incorporation of ecological complexity into the models discussed. I sympathize with both of these criticisms. I agree with the reviewer that allele frequency affects the strength of selection and that the mechanism of selection can be density dependent, especially when soft selection is primary. The literature on soft and hard selection is vast and could likely use a synthesis. However, the objective of this manuscript is to systematically explore natural patterns that may be explained by an underlying process of negative frequency-dependent selection and those that are compatible and possibly more easily explained by other process. As this reviewer points out, I am exploring only a very narrow and specific subset of how allele frequency impacts selection in negative frequency dependent selection. Negative frequency dependent selection is commonly invoked to explain stable polymorphisms even when rare variants are not advantageous because of their rarity (and common variants are not disadvantaged because of their commonness), a fundamental assumption of this framework that must occur. I agree with the reviewer that the ecological interactions that occur that result in a rare variant being advantageous because it is relatively rare can be mediated by density such that both frequency and density matter. However, for negative frequency dependent selection to operate, rare variants must be advantageous because of their relatively rarity, regardless of the ecological interaction that mediates it. I have tried to make this more clear in the manuscript. I have also tried to improve the clarity regarding distinctions between relative and absolute abundance. Lastly, I apologize for writing so much, but I do think these topics are important and explaining them is often not simple.

**In this conceptual paper, the author argues that negative frequency-dependent selection is not necessarily the main mechanism explaining the maintenance of polymorphism.**

The objective of this manuscript is to explore patterns that may be explained by an underlying process of negative frequency-dependent selection and those that are more easily explained by other process. There are many examples of polymorphisms maintained by negative frequency dependent selection that are not controversial.

**The author reviews four other mechanisms that he argues can be mistaken for negative frequency-dependent selection: directional selection in changing environments, density-dependent fitness, multiple niche polymorphisms and community diversity. Although this is an interesting theoretical question, with important empirical implications, I have several issues with the author's treatment. First, many arguments given by the author hinge upon a restrictive and debatable definition of frequency dependence.**

I agree with the reviewer that the central framework explored, along with its assumptions, is a very narrow subsection of the impact of allele frequency on the strength and direction (advantageous, neutral, or disadvantageous) of selection. The central point of this work is not to explore the rather large body of work regarding the interaction between selection and allele frequency, but the rather narrow section of this literature that deals specifically with negative frequency-dependent selection, which is often used as a heuristic explanation for the selective maintenance of genetic variation within populations without much exploration of the fit of the system or data to the framework. I do not know of any debate concerning the definition of negative frequency-dependent selection given in this manuscript, if one exist I would be gratefully if the reviewer would direct me to it. Second, the various examples are only discussed verbally, although the concept of frequency dependence has its roots in mathematical models of population genetics.

**The author defines negative frequency-dependent selection as a biological scenario where "the selective value of an allele is dependent on its relative abundance in the population such that Darwinian fitness increases as the relative abundance, or frequency, of the allele decreases". This definition fits perfectly the original population genetics definition of frequency dependence, but it becomes ambiguous in ecologically realistic populations where, with this literal definition of frequency dependence, selection is always frequency-dependent.**

I agree with the reviewer that additional clarity is needed surrounding the definition of negative frequency dependent selection as well as the purpose and direction of this synthesis. I have rewritten this section to add clarity, specifically referring to the sign change of selection, from advantageous to disadvantageous, as a function of frequency, which is the key assumption of negative frequency dependent selection that was not clearly stated. I do, however, disagree that this definition is ambiguous in ecologically realistic populations. There are many (maybe hundreds) of examples of stable polymorphisms that are unambiguously caused by negative frequency dependent selection, some of which are now described in the manuscript. In these natural systems, there is no ambiguity with regard to the negative frequency dependent selection frameworks being applied. Most importantly for this manuscript, the inaccurate assumption that certain natural polymorphisms are maintained by negative frequency dependent selection, when in reality a different force is maintaining the polymorphism, is not due to a misunderstanding of the ecological complexity but to a misunderstanding of the basic framework. Lastly, I agree that all soft selection is frequency dependent, but that does not indicate that there is always a negative frequency dependent component of all selection nor that all frequency dependent selection, negative or otherwise, results in stable polymorphisms. I have clarified the focus of the manuscript and the definitions throughout.

**For instance, in a density-regulated population at equilibrium, a rare allele will increase in frequency if its per-capita growth rate is positive, but, if it goes to fixation, its per-capita growth rate is necessarily zero because the population is at equilibrium. So we have negative frequency dependence according to the author's definition, but without more ecological details, there is no reason to assume that this will lead to a stable polymorphism.**

This example appears to me to represent directional selection as opposed to negative frequency dependent selection. It would only fit the framework of negative frequency dependent selection if the positive per capita growth rate was the direct result of its rarity. That is, just because an allele increases when rare does not indicate that selection negatively depends on its frequency. I have modified the definition to make more clear that in models of negative frequency dependent selection, a rare variant is advantageous because of its rarity and common variants suffer a disadvantage because of their commonness. I do agree with the reviewer that not all negative frequency dependent selection will lead to a stable polymorphism, either due to ecological or model specifics. I have clarified the focus of the manuscript on the common use of negative frequency dependent selection as a heuristic explanation of stable polymorphisms that are more accurately explained by other mechanisms.

**More generally, fitness is a property of a type in a given environment, so always depends on what others are doing except in very simplistic models. This leads to frequency dependent selection, which is mediated through this environmental feedback. The ecological extension of evolutionary game theory known as adaptive dynamics relies on this notion of frequency dependence.**

I agree with the reviewer that soft selection in both models and nature depends on the frequency of alleles in the population; the environment determines the selective value of an allele. In cases where negative frequency dependent selection is acting to maintain genetic diversity, the relative selective value of an allele changes from advantageous at low frequencies to disadvantageous at high frequencies.

**In addition, the distinction between density- and frequency-dependent selection is also often misleading because frequency-dependence is generally mediated by the densities of different types. Again, the distinction between density- and frequency-dependent selection comes from population genetics models with restrictive ecological assumptions so that it is possible to decouple the impact of total population size from the impact of allele frequencies. This will generally not extend to models of competition, predation or parasitism, for instance.**

Thank you for pointing out this confusion. I was not trying to make a distinction between density- and frequency-dependent selection. I agree that the mechanism affecting survival/reproduction in systems where a polymorphism is maintained by negative frequency dependent selection can be mediated by density-dependent interactions such that the strength of selection is density dependent. Nevertheless, the sign – whether an allele is advantageous, neutral, or disadvantageous – must depend on the frequency of alleles if negative frequency dependent selection is maintaining the polymorphism. The primary point, which I have clarified in this section, is that in some situations there is no evidence that rare variants are advantageous because they are rare, despite the conclusions of these papers, and that the dynamics they described were the result of the densities in the population. I agree with the reviewer that decoupling the density and frequency effects is an advantage of theoretical models. Nevertheless, there are many models and real systems where allele frequency and total density are sufficiently uncoupled and tractable in modes that extend to competition, predation or parasitism (ex. Allen, 1988; Gross, 1991; Harvey et al., 1975; Hori, 1993). Decoupling these effects makes the processes easier to discuss, but the decoupling is not necessary for negative frequency dependent selection to maintain a stable polymorphism.

**The model described by eqn 1, for instance, is used by the author as a classical example of density- but not frequency-dependent fitnesses, but I disagree with this characterisation. First, eqn 1 does depend on the frequencies of each type since it depends on the densities N1 and N2 and not only on the total density N1+N2. Second, these Lotka-Volterra equations have been repeatedly used to showcase the possibility of evolutionary branching where directional selection pushes the population towards a point where selection becomes disruptive and leads to the emergence and maintenance of polymorphism (see e.g. Kisdi JTB 1999). At such a branching point, the population sits at a fitness minimum so any rare mutant can invade (advantage to the rare types). I therefore contest the distinction the author makes about the respective roles of relative vs absolute abundances. I don't think it is meaningful from an ecological perspective. More importantly, it highlights the limitations of the narrow definition of frequency dependence the author uses.**

I agree with the reviewer on several points here. First, I apologize for the over simplistic strawman description here (and throughout). The point of this section was not attempting to make a case for the relative roles of relative vs absolute abundance for selection in general, but to specifically stress that polymorphisms that are maintained by negative frequency dependent selection must demonstrate a negative correlation between relative abundance and relative fitness. These simplified models are used to demonstrate that these models do not produce a sign change (advantageous to disadvantageous) that is correlated with frequency. Although I recognize the confusion, I do disagree with the reviewer that interactions that depend on the density of multiple variants is equivalent to interactions that depend on the abundance of each variant relative to the other variants. For example, many of the dynamics in these models would be lost if N1 and N2 were changed to frequencies p=N1/N and q=N2/N. Most importantly for this piece, the per capita fitness changes occur because of how close the variant is to its carrying capacity (accounting for the competition), not as a function of its relative frequency per se. The paper referenced by the reviewer (Kisdi et al 1999) makes this point as variants with lower competitive abilities (small variants) can invade and grow to high densities because although "small individuals cannot win a contest against large ones, they have a good chance to avoid any such contest altogether, because the large strategy can maintain only a low population density due to the cost of large trait values." It is not the frequency of large variants that result in few interactions with an invading small variant, but the low abundance of the large variant, suggesting that these distinctions may well be ecologically meaningful (see Abrams and Ginzburg, 2000 TREE for another example of the differing impact of frequency and density in ecology). While the paper pointed out by the reviewer is very interesting, I disagree with the reviewer's interpretations of the results, as I understand them. First, this and other papers have been used to understand polymorphisms through asymmetric competition and directional selection, but do not address how negative frequency dependent selection can maintain a polymorphism. These models have not been mistaken for negative frequency dependent selection and thus were not included in the prior submission. I have made a note of these types of models in the revised manuscript as I agree with the reviewer that they are important. Second, while this paper does make the point that a population can get stuck at a branching point (a fitness minimum) due to free random mating resulting in selection against intermediates, it does not suggest that rare variants can invade because they are rare. In fact, free recombination will "prevent the emergence of two phenotypically separate branches."

**An additional complexity comes from the fact that the fitness concepts for short-term vs long-term evolution are different. Frequency-dependence, as defined by the author, is well defined in a short-term context because we can often forget about density-dependence in this case, at least initially. However, for long-term evolution, we need to account for the feedback of the environment, so we need an evolutionary game theory (or frequency-dependent) perspective. This issue should also be clarified.**

I apologize for not completely understanding this comment. Negative frequency dependent selection has a feedback mechanism where the relative selective value of an allele is determined by its frequency in the environment, leading to changes in the frequency which feedback on the selective value. The long-term implications are, under the appropriate conditions (parameter space or ecological context), a stable polymorphism which can be either stationary or cyclic. Thus, both short- and long-term evolutionary dynamics are addressed in these models and in the ecological systems where they are found. I agree with the reviewer that density can be one of the conditions affecting the strength of selection (population size is nearly always a consideration in soft selection), but this is often context specific.

**My own opinion on the matter is that the main conceptual tool to explain the maintenance of polymorphisms is not frequency dependence, which is both an ubiquituous and ambiguous concept, but the dimensionality of the environment. This concept is more general, and encompasses the various examples discussed by the author.**

I agree that many natural polymorphisms are maintained by processes of than negative frequency dependent selection. This paper is meant to bring attention to the error of using of negative frequency dependent selection as a heuristic explanation of stable polymorphisms that are more accurately explained by other mechanisms.

**Other points: • The model analysed in figure 3 is impossible to understand because its details are only given in an R code. The life cycle needs to be given in detail. Furthermore, a mathematical analysis of the model should be possible, given its apparent simplicity, and would be more instructive.**

These models have been published elsewhere and the simulated data were shown to illustrate the points. I have added mathematical details in the revision for clarity.

**• I don't understand the fourth example. It seems the author suddenly looks at the frequency of species instaed of allele frequencies. It is not clear to me how this fits with the original discussion. This needs to be clarified.**

I agree and have changed the introduction to this section to make it more consistent with the rest of the paper. A primary point of the paper is that many authors have been using a negative frequency dependent selection framework to explain diversity that is better explained in a different framework, in this case community diversity.

**Summary:**

**In its current form, I found the paper rather confusing and unconvincing. In part, this is due to the concept of frequency-dependence itself, which is often loosely defined in the literature. The author uses a very literal definition, in line with the original population genetics concept, but, since selection is always frequency-dependent in ecologically realistic models according to that definition, another definition should be used to explain the maintenance of stable polymorphisms. This needs to be clarified, and the arguments backed by a more thorough mathematical analysis, since verbal discussions of numerical simulations only add to the confusion.**

I hope the revision is more clear. **Reviewed by David Baltrus, 2017-05-09 17:31**

**As the "big data revolution" progresses and biology is confronted with ever more complicated patterns to interpret, evolutionary terms are being increasingly invoked to explain perceived patterns. "Frequency dependence" is one of these terms. The purpose of this manuscript from Brisson is to begin to clarify when it is/is not appropriate to use the term "negative frequency dependent selection (NFDS)" in the context of evolutionary explanations. Brisson does a great job of laying out definitions and explanations for use of this term over the last century or so, and does so while describing how such selection regimes could help to explain the amount of diversity we see in the world. I'm a proponent of clearly laying out the case for when nuanced evolutionary terms are applied inappropriately, and Brisson does a good job of describing instances where patterns may suggest negative frequency dependent selection but where this specific evolutionary model doesn't apply. He makes this case throughout the manuscript and does so in a way that is clear and concise. I think this manuscript could go a long way towards clearing up some confusion in the literature if the right people see it at the right time.**

**I have no major qualms with this preprint, it's laid out and written quite well. However, I do think that it would make the case slightly more clear if, in cases where the pattern suggest NFDS falsely, if some examples were imagined that would allow the patterns to fall under the purview of NFDS. For example...what would need to happen to make the "killing the winner" scenario actually fall under**

**NFDS? I'm not sure if there is actually a clear way to do this or if it would muddle things, but if possible it would be good to include additions that could make these situations fall under NFDS as counterpoints.**

This is an interesting point. I have tried to do this with some of the examples, although not the killing the winner example, as I think that natural selection models should not be applied to communities as most of the assumptions of natural selection in general are violated.

I really enjoyed this preprint both for its subject matter and clarity, and I hope to see it well received across communities.

## Decision by Ignacio Bravo, posted 08 May 2017

**Revision needed**

Two experts in the field have reviewed the preprint entitled ”Negative frequency-dependent selection is frequently confounding”. Although the reviewers' impressions on the text are largely divergent, I consider that they remain compatible, and I subscribe (to different extents) to their opinions. Thus, I profoundly agree with the author in the need of a systematic identification of patterns that may be explained by an underlaying process of negative frequence-dependent selection (NFDS), but that may also be compatible and possibly more easily explained by different process. I also acknowledge the effort in referring to the original literature (although in the case of referring to books, I would appreciate if the quote could be identified with more precision, including pages or at least chapters).

The author elaborates on four processes that may lead to maintenance of diversity and whose results may be mistaken for NFDS: directional selection in changing environments, density-dependent fitness, multi-niche selection and community diversity. Reviewer#2 has raised a number of important points in the direction of how the width of the definition used for NFDS has an impact on the interpretation/usefulness of the concept itself. While the author has chosen to stick to a classical population genetics definition of NFDS, the reviewer claims that this framework may not be appropriate for analysing the ecological and long-term implications of a genuine NFDS. These conflicting views need to be reconciled. The reviewer criticises then the difference used by the author between relative and absolute abundances, with respect to the carrying capacity of the environment, and its implication. I think that this is a major criticism that also requires to be addressed. The same applies to the appropriateness of the verbal arguments, the completeness of the example used in Figure 3 and the lack of a thorough mathematical analysis. Finally, I also agree with both reviewers that the fourth process “community diversity” is the less-well developed and somehow disconnected with the discussion of allele/variant frequencies in the three previous processes.

Globally, I consider that the aim of the review is extremely pertinent and that it may be very useful to clarify when NFDS is truly the most likely explanatory process for a given observed pattern. Nevertheless, a number of important points need to be addressed, and the text and largely benefit from a more explicit introduction of ecology into the population genetics definition of NFDS and from the mathematical and not only verbose systematisation of different patterns being compatible with different underlying processes.

## Reviewed by anonymous reviewer 1, 17 May 2017

In this conceptual paper, the author argues that negative frequency-dependent selection is not necessarily the main mechanism explaining the maintenance of polymorphism. The author reviews four other mechanisms that he argues can be mistaken for negative frequency-dependent selection: directional selection in changing environments, density-dependent fitness, multiple niche polymorphisms and community diversity.

Although this is an interesting theoretical question, with important empirical implications, I have several issues with the author's treatment. First, many arguments given by the author hinge upon a restrictive and debatable definition of frequency dependence. Second, the various examples are only discussed verbally, although the concept of frequency dependence has its roots in mathematical models of population genetics.

The author defines negative frequency-dependent selection as a biological scenario where ”the selective value of an allele is dependent on its relative abundance in the population such that Darwinian fitness increases as the relative abundance, or frequency, of the allele decreases”. This definition fits perfectly the original population genetics definition of frequency dependence, but it becomes ambiguous in ecologically realistic populations where, with this literal definition of frequency dependence, selection is always frequency-dependent. For instance, in a density-regulated population at equilibrium, a rare allele will increase in frequency if its per-capita growth rate is positive, but, if it goes to fixation, its per-capita growth rate is necessarily zero because the population is at equilibrium. So we have negative frequency dependence according to the author’s definition, but without more ecological details, there is no reason to assume that this will lead to a stable polymorphism. More generally, fitness is a property of a type in a given environment, so always depends on what others are doing except in very simplistic models. This leads to frequency dependent selection, which is mediated through this environmental feedback. The ecological extension of evolutionary game theory known as adaptive dynamics relies on this notion of frequency dependence.

In addition, the distinction between density- and frequency-dependent selection is also often misleading because frequency-dependence is generally mediated by the densities of different types. Again, the distinction between density- and frequency-dependent selection comes from population genetics models with restrictive ecological assumptions so that it is possible to decouple the impact of *total* population size from the impact of allele frequencies. This will generally not extend to models of competition, predation or parasitism, for instance. The model described by eqn 1, for instance, is used by the author as a classical example of density- but not frequency-dependent fitnesses, but I disagree with this characterisation. First, eqn 1 does depend on the frequencies of each type since it depends on the densities N1 and N2 and not only on the total density N1+N2. Second, these Lotka-Volterra equations have been repeatedly used to showcase the possibility of evolutionary branching where directional selection pushes the population towards a point where selection becomes disruptive and leads to the emergence and maintenance of polymorphism (see e.g. Kisdi JTB 1999). At such a branching point, the population sits at a fitness minimum so *any rare mutant* can invade (advantage to the rare types). I therefore contest the distinction the author makes about the respective roles of relative vs absolute abundances. I don’t think it is meaningful from an ecological perspective. More importantly, it highlights the limitations of the narrow definition of frequency dependence the author uses.

An additional complexity comes from the fact that the fitness concepts for short-term vs long-term evolution are different. Frequency-dependence, as defined by the author, is well defined in a short-term context because we can often forget about density-dependence in this case, at least initially. However, for long-term evolution, we need to account for the feedback of the environment, so we need an evolutionary game theory (or frequency-dependent) perspective. This issue should also be clarified.

My own opinion on the matter is that the main conceptual tool to explain the maintenance of polymorphisms is not frequency dependence, which is both an ubiquituous and ambiguous concept, but the dimensionality of the environment. This concept is more general, and encompasses the various examples discussed by the author.

Other points:

- The model analysed in figure 3 is impossible to understand because its details are only given in an R code. The life cycle needs to be given in detail. Furthermore, a mathematical analysis of the model should be possible, given its apparent simplicity, and would be more instructive.
- I don’t understand the fourth example. It seems the author suddenly looks at the frequency of species instaed of allele frequencies. It is not clear to me how this fits with the original discussion. This needs to be clarified.

Summary:

In its current form, I found the paper rather confusing and unconvincing. In part, this is due to the concept of frequency-dependence itself, which is often loosely defined in the literature. The author uses a very literal definition, in line with the original population genetics concept, but, since selection is always frequency-dependent in ecologically realistic models according to that definition, another definition should be used to explain the maintenance of stable polymorphisms. This needs to be clarified, and the arguments backed by a more thorough mathematical analysis, since verbal discussions of numerical simulations only add to the confusion.

## Reviewed by David Baltrus, 17 May 2017

As the ”big data revolution” progresses and biology is confronted with ever more complicated patterns to interpret, evolutionary terms are being increasingly invoked to explain perceived patterns. ”Frequency dependence” is one of these terms. The purpose of this manuscript from Brisson is to begin to clarify when it is/is not appropriate to use the term ”negative frequency dependent selection (NFDS)” in the context of evolutionary explanations. Brisson does a great job of laying out definitions and explanations for use of this term over the last century or so, and does so while describing how such selection regimes could help to explain the amount of diversity we see in the world. I'm a proponent of clearly laying out the case for when nuanced evolutionary terms are applied inappropriately, and Brisson does a good job of describing instances where patterns may suggest negative frequency dependent selection but where this specific evolutionary model doesn't apply. He makes this case throughout the manuscript and does so in a way that is clear and concise. I think this manuscript could go a long way towards clearing up some confusion in the literature if the right people see it at the right time.

I have no major qualms with this preprint, it's laid out and written quite well. However, I do think that it would make the case slightly more clear if, in cases where the pattern suggest NFDS falsely, if some examples were imagined that would allow the patterns to fall under the purview of NFDS. For example...what would need to happen to make the ”killing the winner” scenario actually fall under NFDS? I'm not sure if there is actually a clear way to do this or if it would muddle things, but if possible it would be good to include additions that could make these situations fall under NFDS as counterpoints.

I really enjoyed this preprint both for its subject matter and clarity, and I hope to see it well received across communities.