HAL
open science

Fisher to the rescue

François Blanquart, Florence Débarre

To cite this version:

François Blanquart, Florence Débarre. Fisher to the rescue. 2019, pp.100072. 10.24072/pci.evolbiol.100072 . hal-02550744

HAL Id: hal-02550744

https://hal.science/hal-02550744v1

Submitted on 29 Apr 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Peer Community In
Evolutionary Biology

Fisher to the rescue

François Blanquart and Florence Débarre based on reviews by 3 anonymous reviewers

A recommendation of:

Florian Lavigne, Guillaume Martin, Yoann Anciaux, Julien Papaïx, Lionel Roques. **When sinks become sources: adaptive colonization in asexuals** *(2019), bioRxiv, 433235, ver. 5 peer-reviewed and recommended by Peer Community in Evolutionary Biology.* 10.1101/433235

Open Access

Submitted: 03 October 2018, Recommended: 30 April 2019

Cite this recommendation as:
François Blanquart and Florence Débarre (2019) Fisher to the rescue. *Peer Community in Evolutionary Biology, 100072.* 10.24072/pci.evolbiol.100072

Published: 6 May 2019

Copyright: This work is licensed under the Creative Commons Attribution-NoDerivatives 4.0 International License. To view a copy of this license, visit http://creativecommons.org/licenses/by-nd/4.0/

The ability of a population to adapt to a new niche is an important phenomenon in evolutionary biology. The colonisation of a new volcanic island by plant species; the colonisation of a host treated by antibiotics by a-resistant strain; the Ebola virus transmitting from bats to humans and spreading epidemically in Western Africa, are all examples of a population invading a new niche, adapting and eventually establishing in this new environment.

Adaptation to a new niche can be studied using source-sink models. In the original environment —the "source"—, the population enjoys a positive growth-rate and is self-sustaining, while in the new environment —the "sink"— the population has a negative growth rate and is able to sustain only by the continuous influx of migrants from the source. Understanding the dynamics of adaptation to the sink environment is challenging from a theoretical standpoint, because it requires modelling the demography of the sink as well as the transient dynamics of adaptation. Moreover, local selection in the sink and immigration

from the source create distributions of genotypes that complicate the use of many common mathematical approaches.

In their paper, Lavigne et al. [1], develop a new deterministic model of adaptation to a harsh sink environment in an asexual species. The fitness of an individual is maximal when a number of phenotypes are tuned to an optimal value, and declines monotonously as phenotypes are further away from this optimum. This model —called Fisher's Geometric Model— generates a GxE interaction for fitness because the phenotypic optimum in the sink environment is distinct from that in the source environment [2]. The authors circumvent mathematical difficulties by developing an original approach based on tracking the deterministic dynamics of the cumulant generating function of the fitness distribution in the sink. They derive a number of important results on the dynamics of adaptation to the sink:

- From the point where immigration from the source to the sink starts, four phases of adaptation are observed. After a short transient phase (phase 1), a migration-selection balance is reached in the sink (phase 2). After a while, thanks to the immigration of rare adapted migrants and mutation in the sink, a small fraction of the sink population exhibits a close-to-optimal phenotype. This small adapted fraction grows in frequency and mean fitness rapidly increases in the sink (phase 3). Finally, the population settles around the sink optimum (phase 4) and, hurray, the sink is now a source!

- Interestingly, in this model the evolutionary dynamics do not depend on the immigration rate. In other words, adaptation will proceed at the same rate regardless of how many immigrants invade the sink. This is because the impact of immigration on adaptation depends on the rate of immigration relative to the sink density. This ratio is actually independent of immigration in a model where the sink is initially empty, migration from the sink back to the source is negligible and without density-dependence in the sink.

- In this model, mutation is a double-edged sword. Adapted phenotypes emerge from new mutations, and under this effect alone a higher mutation rate would translate into a shorter time to establishment in the sink. However, mutations may also have deleterious effects by displacing the phenotype away from the optimum. This mutation load will be greater when individuals need to simultaneously tune a large number of phenotypes. As a consequence of these two effects of mutations, time to establishment is

minimal for an intermediate mutation rate. This result emerges from Fisher's Geometric Model, but may hold more generally for biologically plausible fitness landscapes where mutations generates both beneficial (allowing adaptation to the sink) and deleterious genotypes.

- Lastly, in Fisher's Geometric Model, the time to establishment increases superlinearly with harshness of the sink when the sink is too harsh, and establishment may occur only after a very long time. In these harsh sinks, the adapted genotypes are very few and increase very slowly in frequency, making the second phase of adaptation much longer. Thus, and as a direct consequence of Fisher's Geometric Model, adding a "stepping stone" intermediate environment would allow faster adaptation to the extreme environment.

In conclusion, this theoretical work presents a method based on Fisher's Geometric Model and the use of cumulant generating functions to resolve some aspects of adaptation to a sink environment. It generates a number of theoretical predictions for the adaptive colonisation of a sink by an asexual species with some standing genetic variation. It will be a fascinating task to examine whether these predictions hold in experimental evolution systems: will we observe the four phases of the dynamics of mean fitness in the sink environment? Will the rate of adaptation indeed be independent of the immigration rate? Is there an optimal rate of mutation for adaptation to the sink? Such critical tests of the theory will greatly improve our understanding of adaptation to novel environments.

References

[1] Lavigne, F., Martin, G., Anciaux, Y., Papaïx, J., and Roques, L. (2019). When sinks become sources: adaptive colonization in asexuals. bioRxiv, 433235, ver. 5 peer-reviewed and recommended by PCI Evolutionary Biology. doi: 10.1101/433235

[2] Martin, G., and Lenormand, T. (2006). A general multivariate extension of Fisher's geometrical model and the distribution of mutation fitness effects across species. Evolution, 60, 893-907. doi: 10.1111/j.0014-3820.2006.tb01169.x

Revision round #2

2019-04-29

The editors and reviewer think the authors have satisfactorily modified the manuscript. The new paragraph "Phenotypic dynamics over the different phases of invasion", line 312, is particularly helpful.

Just a few minor comments:

line 81 I would remove the first "resp."

line 188 and 190: please change dz to dx for consistency with the new notation (e.g. line 272).

line 415 "sustained by" -> "supported by".

Figure 6 legend: please explain what the -dU/rbar(0) = 500 corresponds to.

We are currently working on the recommendation text.

Additional requirements of the managing board: As indicated in the 'How does it work?' section and in the code of conduct, please make sure that: -Data are available to readers, either in the text or through an open data repository such as Zenodo (free), Dryad (to pay) or some other institutional repository. Data must be reusable, thus metadata or accompanying text must carefully describe the data. -Details on quantitative analyses (e.g., data treatment and statistical scripts in R, bioinformatic pipeline scripts, etc.) and details concerning simulations (scripts, codes) are available to readers in the text, as appendices, or through an open data repository, such as Zenodo, Dryad or some other institutional repository. The scripts or codes must be carefully described so that they can be reused. -Details on experimental procedures are available to readers in the text or as appendices. -Authors have no financial conflict of interest relating to the article. The article must contain a "Conflict of interest disclosure" paragraph before the reference section containing this sentence: "The authors of this preprint declare that they have no financial conflict of interest with the content of this article." If appropriate, this disclosure may be completed by a sentence indicating that some of the authors are PCI recommenders: "XXX is one of the PCI XXX recommenders."

Preprint DOI: 10.1101/433235

Reviewed by anonymous reviewer, 2019-04-12 17:49

I am quite happy with the replies that the authors have given to the point raised during the first round. I have checked the modifications - in particular, I was able to double-check that the proofs given in the appendix are understandable and

correct. I therefore recommend the article to be accepted for the PCI Evol Biol recommendation.

Author's reply:

We thank the recommenders and the reviewers for their impressive work. We have corrected all the minor points in the revised version of our MS : https://www.biorxiv.org/content/10.1101/433235v4

Revision round #1

2018-12-14

First, I would like to apologize to the authors for the time it has taken to evaluate their preprint. However, I was lucky to eventually find three specialists who agreed to spend time evaluating the preprint -- many thanks to them!

The authors will notice that the evaluators have different backgrounds, and that their opinions about the manuscript differ a bit. All of their comments are useful and I encourage the authors to take them into account while revising their manuscript.

In particular, I share a reviewer's comment about the lack of biological interpretation of the results. In its current version, the manuscript is very much focused on solving equations, obtaining analytical expressions, rather than providing biological interpretations in the context of the model. I feel that the manuscript is currently rather written for a mathematical biology audience, than for an evolutionary biology audience. It is really just a matter of presentation; I do think that the results can be of interest for evolutionary biologists.

In addition to the reviewer's comments, I have a few remarks myself, presented as they appear in the manuscript.

l.42-ff Could be rephrased: as currently written, we expect a description of other types of source-sink systems as well.

l.71-125: I found these two paragraph quite hard to follow; the actual aim of these paragraphs is not clear.

l.84-ff I do not understand this description of “fitness-based” (the “selection on fitness itself” part; fitness is also to be taken into account in the “trait-based” versions, no?)

l.107 I disagree: gene swamping does not only occur in sexuals. See for instance Nagylaki 1975 (PMC1213362).

l.114 Same comment: gene swamping does also occur in asexual populations.

l.152-153, 167-168... Unnumbered lines. This is a well-known bug of the lineno package, and there are fixes to ensure that all lines are correctly numbered (e.g. http://phaseportrait.blogspot.com/2007/08/lineno-and-amsmath-compatibility.html)

unnumbered line above unnumbered equation above l.153: This should rather read “mean absolute and mean relative fitness”. Also, it is rather odd to call m a “relative fitness”; usually, the term is used for a fitness divided by the mean(/total) fitness in the population. Please consider using another term (e.g. fitness differential?)

unnumbered equation above l.153 Please define notation $\bar$

equation (1) Please define notation $’$ (prime). (sometimes used in population genetics for different equations).

l.235-236 Moment k usually depends not just on moment k+1.

Fig.1 $\bar{r}(0)$ could be identified on the horizontal axis and its value given in the legend. Similarly, the value 0 could be better highlighted.

l.337 “migrants do not further breed with and genetically “pollute” locally adapted genotypes.” I think that this sentence should be re-written to avoid being misused.

l.383ff The LaTeX command for properly typeset “>>” is $\gg$.

l.397ff (and 511-ff) The discussion on the effect of an intermediate sink is interesting, but it would be more impactful if it came with a corresponding figure (and model).

l.440 There seems to be a problem at the end of the line (missing words).

Preprint DOI: 10.1101/433235

Reviewed by anonymous reviewer, 2018-10-08 15:58

The authors analyse a model of a population adapting to a harsh sink environment under sustained migration from a source. The fitness landscape is defined by Fisher's Geometric Model, where the population needs to simultaneously tune a number of phenotypes to their optimal values. They derive a number of analytical results on the dynamics of adaptation under a deterministic approximation that holds when the mutation rate and the migration rate are high.

The main results are the following: 1. The model predicts that successful adaptation to the sink proceeds in three phases: after a short initial phase of rapid increase in fitness, the fitness plateaus for some time then re-accelerates and reaches its equilibrium values 2. The dynamics of adaptation does not depend on the immigration rate 3. When the population adapts to the sink, the migration load is negligible because the sink population becomes very large 4. Above a certain mutation rate, the population may not adapt to the sink. It then suffers from a migration load. 5. Adaptation proceeds faster (lower establishment time) when the mutation rate, the mutational variance, and the maximal growth rate are higher. 6. In harsh environments, the time to establishment increases faster than linearly with the harshness of stress. In those cases, intermediate habitats could greatly speed up adaptation.

Generally, I enjoyed reading this interesting manuscript. However, as a biologist, I would have liked the analytical results to be interpreted in biological terms more often. This is particularly critical for some of the main results that are not biologically interpreted: the existence of three phases, the fact that adaptation does not depend on the immigration rate, the fact that the time to establishment increases faster than linearly with harshness of the sink. It is difficult to assess whether these results emerge generally from source-sink dynamics, are properties of the specific patterns of epistasis or GxE imposed by Fisher's Geometric Model, or of the specific high migration / high mutation regime investigated here. It is important that the authors delineate better the range of applicability of these results.

Along the same lines of conveying an intuition for when the results work, it would be interesting to develop the types of biological systems where these results could apply. The authors mention pathogens. Indeed many pathogens are clonal and present high diversity, but the transmission rate (immigration rate) is not necesarily high, so within-host diversity may not be as high as that investigated by

the authors. Maybe (just speculating here!) bacteria in water environments would correspond to that regime.

Lastly, it would be interesting to draw a figure with the evolving distribution of fitness in the sink. I am imagining that it initially resembles the distribution shown on figure 1, then it progressively shifts to the right until it reaches a positive mean growth rate. Such as figure would also show why classical mathematical approaches (for example constant variance) fail.

Major comments along these lines:

line 251 Some time could be devoted to explaining the different terms in equation (9). Most evolutionary biologists, including theoreticians, do not think of evolutionary dynamics in terms of cumulant generating function. Why is the selection term depending on the difference with d*zC*t(0)? What is the (n/2) * z term in mutation? Is the effect of migration to make C_t(z) ressemble phi(z) ?

line 290 those three phases are one the most important result, yet the authors do not explain what phenomena occur in these three phases. What is the initial brief increase in fitness followed by a plateau ? Why is fitness accelerating again? These seems to depend on harshness of the environment m_D (line 291) but we do not have more information. Is this due to the curvature of the FGM? Or is it a more general phenomenon relevant to a larger class of models of adaptation to a sink environment?

line 305 the authors could add an interpretation for the interesting lack of effect of d on evolutionary dynamics. Essentially (if my understanding is correct) this is because the local population size in the sink is directly proportional to the influx of migrants d, so the fraction of migrants in the sink is constant regardless of d. This effect is expected to be true for any source-black hole sink dynamics, is there mention of this in previous literature ? Maybe the Gomulkiewicz et al 1999 ?

line 346 the duration of "phase 1" is more variable when the mutation rate mu is decreased. But again it would be interesting to have an interpretation of this phase 1, and what phenomenon occurs when phase 2 starts.

line 398 any intuition for why there is such a threshold for the harshness of stress beyond which adaptation does not occur? What features of the model does it depend on?

Other comments:

line 13-14 "stress induced by migration" is misleading. Maybe "stress induced by the sink environment"? likewise lines 150

line 24 phaseS

line 54 The conjunction "As such" does not connect well with the paragraph above

line 80 "no model, be it mathematical or simulation-based, can tackle all these various factors together" Maybe specify, e.g. "a model with all these various factors would be too complex to be intelligible" if this is what the authors mean.

line 91 I think that epistasis and GxE interactions for fitness can be incorporated into a fitness-based model. Perhaps say something more specific, like the distribution of epistasis and GxE interaction generated by FGM match empirical data well.

line 122-123 the infinitesimal model makes assumptions on the genetic architecture of traits. These assumptions can be used for clonal species as well as sexual ones. Likewise, is the constant genetic variance assumption always breaking down in asexuals? I would say it might work when the mutation rate is high compared to selection, even in clonal species.

Table 1 the use of z as the variable of the cumulant generating function is confusing as this is also the name of the phenotype vector. Use x?

after line 167, "in average" -> "on average"

line 192 rephrase 'the decay rate, in the sink, of an optimal phenotype'. This is the decay rate of a population composed of individuals with the optimal phenotype only.

line 202 could you please comment on the distribution of effects of mutations in FGM? For example say that its mean is -lambda / 2 * n, also specify its variance

line 215 please define the notation Gamma. It would be nice to interpret this distribution further. What is the mean fitness, the variance of fitness

Below equation 7, it would be nice to have an interpretation of the -mu * n/2 term (mutation load). And a comment on the fact that it depends on dimensionality (is it the so-called 'cost of complexity'?)

figure 1 would be helpful to place the r_max on the x-axis

line 300 at this point it was not clear why large d is needed.

equation (11) perhaps the result on the asymptotic value of rbar can be compared with the result on the equilibrium m_source (line 215), that has the same mean.

line 370 if I understand correctly, higher rmax means faster adaptation both because it slows the decay of the sink population and increases the proportion of migrants that have positive growth in the sink. The sentence gives the impression that it is not fully resolved why increasing rmax allows faster adaptation, perhaps rephrase.

line 374 would be helpful to refer to figure 1 here.

figure 6, could you please approximately place the m_D beyond which adaptation in the sink does not occur for these parameters?

line 434 again would be nice to have a biological interpretation of these three phases

line 497 would be a good place to refer to fig. 4.

line 533 anisotropic version, very well, but I guess in that case the analysis becomes MUCH harder because the fitness effect of mutation depends not only on the fitness of the parental genotype but also on the exact position in the phenotypic space (the z_i).

Reviewed by anonymous reviewer, 2018-11-08 11:42

Download the review (PDF file)

Reviewed by anonymous reviewer, 2018-11-15 16:03

Download the review (PDF file)

Author's reply:

Download author's reply (PDF file)