



Pseudo-overdominance: how linkage and selection can interact and oppose to purging of deleterious mutations.

Sylvain Glémin based on reviews by Yaniv Brandvain, Lei Zhao and 1 anonymous reviewer

A recommendation of:

Conditions for maintaining and eroding pseudo-overdominance and its contribution to inbreeding depression

Diala Abu Awad, Donald Waller (2022) *bioRxiv*, 2021.12.16.473022, ver. 3 peer-reviewed and recommended by Peer Community in Evol Biol
<https://doi.org/10.1101/2021.12.16.473022>

Open Access

Codes used in this study

- <https://doi.org/10.5281/zenodo.6874326>

Scripts used to obtain or analyze results

- <https://doi.org/10.5281/zenodo.6874326>

Published: 2022-12-16

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/4.0/>

Submission: posted 04 January 2022

Recommendation: posted 15 December 2022, validated 16 December 2022

Cite this recommendation as:

Sylvain Glémin (2022) Pseudo-overdominance: how linkage and selection can interact and oppose to purging of deleterious mutations.. *Peer Community in Evolutionary Biology*, 100531. <https://doi.org/10.24072/pci.evolbiol.100531>

Recommendation

Most mutations affecting fitness are deleterious and they have many evolutionary consequences. The dynamics and consequences of deleterious mutations are a long-standing question in evolutionary biology and a strong theoretical background has already been developed, for example, to predict the mutation load, inbreeding depression or background selection. One of the classical results is that inbreeding helps purge partially recessive deleterious mutations by exposing them to selection in homozygotes. However, this mainly results from single-locus considerations. When interactions among several, more or less linked, deleterious mutations are taken into account, peculiar dynamics can emerge. One of them, called pseudo-overdominance (POD), corresponds to the maintenance in a population of two (or more) haplotype blocks composed of several recessive deleterious mutations in repulsion that mimics overdominance. Indeed, homozygote individuals for one of the haplotype blocks expose many deleterious mutations to selection whereas they are reciprocally masked in heterozygotes, leading to higher fitness of heterozygotes compared to both homozygotes. A related process, called associative overdominance (AOD) is the effect of such deleterious alleles in repulsion on the linked neutral variation that can be increased by AOD. Although this possibility has been recognized for a long time (Otha and Kimura

1969), it has been mainly considered an anecdotal process. Recently, both theoretical (Zhao and Charlesworth 2016) and genomic analyses (Gilbert et al. 2020) have renewed interest in such a process, suggesting that it could be important in weakly recombining regions of a genome. Donald Waller (2021) - one of the co-authors of the current work - also recently proposed that POD could be quantitatively important with broad implications, and could resolve some unexplained observations such as the maintenance of inbreeding depression in highly selfing species. Yet, a proper theoretical framework analysing the effect of inbreeding on POD was lacking.

In this theoretical work, Diala Abu Awad and Donald Waller (2022) addressed this question through an elegant combination of analytical predictions and intensive multilocus simulations. They determined the conditions under which POD can be maintained and how long it could resist erosion by recombination, which removes the negative association between deleterious alleles (repulsion) at the core of the mechanism. They showed that under tight linkage, POD regions can persist for a long time and generate substantial segregating load and inbreeding depression, even under inbreeding, so opposing (for a while) to the purging effect. They also showed that background selection can affect the genomic structure of POD regions by rapidly erasing weak POD regions but maintaining strong POD regions (i.e with many tightly linked deleterious alleles).

These results have several implications. They can explain the maintenance of inbreeding depression despite inbreeding (as anticipated by Waller 2021), which has implications for the evolution of mating systems. If POD can hardly emerge under high selfing, it can persist from an outcrossing ancestor long after the transition towards a higher selfing rate and could explain the maintenance of mixed mating systems (which is possible with true overdominance, see Uyenoyama and Waller 1991). The results also have implications for genomic analyses, pointing to regions of low or no recombination where POD could be maintained, generating both higher diversity and heterozygosity than expected and variance in fitness. As structural variations are likely widespread in genomes with possible effects on suppressing recombination (Mérot et al. 2020), POD regions should be checked more carefully in genomic analyses (see also Gilbert et al. 2020).

Overall, this work should stimulate new theoretical and empirical studies, especially to assess how quantitatively strong and widespread POD can be. It also stresses the importance of properly considering genetic linkage genome-wide, and so the role of recombination landscapes in determining patterns of diversity and fitness effects.

References

Awad DA, Waller D (2022) Conditions for maintaining and eroding pseudo-overdominance and its contribution to inbreeding depression. *bioRxiv*, 2021.12.16.473022, ver. 3 peer-reviewed and recommended by Peer Community in Evolutionary Biology. <https://doi.org/10.1101/2021.12.16.473022>

Gilbert KJ, Pouyet F, Excoffier L, Peischl S (2020) Transition from Background Selection to Associative Overdominance Promotes Diversity in Regions of Low Recombination. *Current Biology*, 30, 101-107.e3. <https://doi.org/10.1016/j.cub.2019.11.063>

Mérot C, Oomen RA, Tigano A, Wellenreuther M (2020) A Roadmap for Understanding the Evolutionary Significance of Structural Genomic Variation. *Trends in Ecology & Evolution*, 35, 561–572. <https://doi.org/10.1016/j.tree.2020.03.002>

Ohta T, Kimura M (1969) Linkage disequilibrium at steady state determined by random genetic drift and recurrent mutation. *Genetics*, 63, 229–238. <https://doi.org/10.1093/genetics/63.1.229>

Uyenoyama MK, Waller DM (1991) Coevolution of self-fertilization and inbreeding depression II. Symmetric overdominance in viability. *Theoretical Population Biology*, 40, 47–77. [https://doi.org/10.1016/0040-5809\(91\)90046-I](https://doi.org/10.1016/0040-5809(91)90046-I)

Waller DM (2021) Addressing Darwin's dilemma: Can pseudo-overdominance explain persistent inbreeding depression and load? *Evolution*, 75, 779–793. <https://doi.org/10.1111/evo.14189>

Zhao L, Charlesworth B (2016) Resolving the Conflict Between Associative Overdominance and Background Selection. *Genetics*, 203, 1315–1334. <https://doi.org/10.1534/genetics.116.188912>

Conflict of interest:

The recommender in charge of the evaluation of the article and the reviewers declared that they have no conflict of interest (as defined in [the code of conduct of PCI](#)) with the authors or with the content of the article.

Reviews

Toggle reviews

Evaluation round #2

DOI or URL of the preprint: <https://doi.org/10.1101/2021.12.16.473022>

Version of the preprint: 2

Author's Reply, 24 Nov 2022

[Download author's reply](#)[Download tracked changes file](#)

Decision by *Sylvain Glémin*, posted 07 Nov 2022, validated 10 Nov 2022

Recommendation on the second version of the manuscript entitled “The role of pseudo-overdominance in maintaining inbreeding depression” by D. Abu-Awad and D. Waller.

First of all, I would like to deeply apologize for the very long delay in answering to this review, which was due to the difficulty of obtaining reviews for this second version and lack of availability from my own side.

We finally obtained two reviews for this revised version. Both are positive and stress the improvement compared to the previous version and the great interest of the results. They only suggested a few minor corrections that can easily be handled. Reviewer 2 noted that the title doesn't fully reflect the main findings of the manuscript. I think it's a matter of taste that can be left to the authors. As the title focus on the maintenance of inbreeding depression, I think it is quite directly related to how long it takes to erode POD regions.

The authors made a great effort to respond to the comments, in particular to clarify the objectives of the study and to extend the analysis of the model to a much broader range of parameters. Overall, it strongly reinforced the initial findings and it should stimulate more empirical work to assess the qualitative and quantitative importance of such POD genomic regions.

I thus consider that this manuscript can be recommended by PCI, but I also recommend the authors to address before the few minor comments raised by the two reviewers (and see the few additional ones below).

Sylvain Glémin

Additional minor comments

L 106: “We assume complete linkage among matched sets of mildly...” I think it should be “ We assume initial complete linkage...” as it can then be broken by recombination as loci are spread every 1 cM.

L 127 “Inbreeding depression δ is a local variable”. What do you exactly mean by “local variable”? Do you mean “within population”?

L 144. The full expression of $I = s_1 s_2 / (s_1 + s_2)$ could be given in the main text and when $s_1 = s_2 = s_H$ it could be noted that $I = s_H / 2$ instead of $s_H^2 / 2 s_H$.

L 144: a space is missing after “Eq. A2”.

L 211 : “At higher mutation rates, singletons will be frequent.” This is not clear to me. The number of singletons should indeed increase but as all kinds of mutations. At stationary state, the proportion of singleton should be the same. Is it correct?

L 365: Note that small genome does not necessarily imply tight linkage because the number of crossovers per chromosome varies little such that recombination rate is higher in small chromosomes. So genes are physically closer on a small chromosome but not necessary more genetically linked.

L 393: “we face the question of what force perpetuates these even within small and inbred populations”

→ add a comma between “these” and “even”

L504: space missing after “viability”

Reviewed by *Lei Zhao*, 26 Oct 2022

The paper has been modified and improved compared with the last version. The authors applied both theoretical and numerical methods to discuss the maintenance and the break-down of the POD (pseudo-overdominant) region. The factors such as spatial distribution of individual mutants, individual selection strength, recombination, dominance, and differentiated fitness between homozygotes, have been taken into consideration. The authors also learned the interactions between the POD and the background mutations occurring elsewhere. I think the manuscript is of great interest and could trigger further studies on the relevant topics.

Some mistakes and concerns are as follows:

1. Line 99, “ $s_x = \min_{s_1, s_2}$ ” should be changed to “ $s_x = \min\{s_1, s_2\}$ ”;
2. Line 144, “see Eq. A2 from Supp. File 1” should be modified as “see Eq. A2 from Supp. File 1”.
3. Line 162, I think what the authors really meant was the “Ordinary Difference Equations” rather than the “Ordinary Differential Equations”, as the equations (10) do not contain the derivatives.
4. Line 192, “The probability that a recombination event occurs between two trans-mutations is then I .” should be modified as “The expected number of recombination events occurring between two trans-mutations is then I .”
5. Figure 3 a and b, I think this might be something wrong, Since the proportion of fixation and loss at some positions are negative (especially for the blue and green solid lines), I cannot see a reason for this.

Reviewed by anonymous reviewer, 02 Nov 2022

Dear editor,

I have reviewed the manuscript entitled “The role of pseudo-overdominance in maintaining inbreeding depression” by Abu Awad and Waller for PCI Evol Biol. I was not involved in the first round of review, and have read the updated version of the MS as well as the comments and replies from the first round.

I really enjoyed reading the MS. Pseudo-overdominance is an important issue in evolutionary biology, and such theoretical models on how pseudo-overdominance zones erode over generations is, I think, a timely contribution. The MS reads well and the findings are interesting. It seems to me that the authors have well addressed the previous comments, such as a better justification for the origin of the POD zone and a better clarity of the results section and figures. I have thus only a few very minor comments.

MINOR COMMENTS

Title: To me, the main message of the MS is about how POD zones erode over time, which is not captured by the title.

line 70: The authors may want to indicate the estimated divergence time between *Capsella* and *Arabidopsis*.

lines 126-156: Having never worked on inbreeding depression, it is unclear to me what 'delta s' and 'delta od' capture. It may need to be explained.

line 222: The verb is missing in "The general map length $R = 1$ and 10 Morgans ...".

lines 258-261: This result may be briefly explained. It is not intuitive to me why, when no mutations have been cleaved off, the recombinant haplotype is selected against. It has the same number of deleterious mutations than the two other haplotype.

Figure 2: In the legend, is sH equal to 0.455, or to 0.45?

Figure S2: The fifth line of the legend seems to be encapsulated in the formula.

Evaluation round #1

DOI or URL of the preprint: <https://doi.org/10.1101/2021.12.16.473022>

Version of the preprint: 1

Author's Reply, 23 Jul 2022

[Download author's reply](#)[Download tracked changes file](#)

Decision by Sylvain Glémin, posted 25 Feb 2022

Recommendation on the manuscript entitled "The role of pseudo-overdominance in maintaining inbreeding depression" by D. Abu-Awad and D. Waller.

This manuscript presents a theoretical model based on both analytical predictions and simulations that addresses the question of the maintenance of pseudo-overdominance and its possible role in explaining inbreeding depression, especially persistence of inbreeding depression in highly selfing species. As myself, the two reviewers appreciated this work, which could be of general interest for the community. However, reviewer 2 raised several concerns about the modelling assumptions, hence the biological conclusions that could be drawn from the results. Both reviewers also asked for clarification in presentation and terminology.

In line with reviewer 2, I think the two main weaknesses of the manuscript are: i) the quite loose hypothesis about the origin of POD regions, and ii) the limited exploration of deleterious mutations parameters.

About the plausibility of the scenario for the origin of POD, if I understood correctly, the initial heterosis experienced by the hybrid is very high: the hybrid as a fitness twice as high as the parental population $1/(1-sH)$. This seems rather unlikely.

The numerical example used is based on some empirical data but is somewhat arbitrary. It shows that many weakly, partially recessive mutations are required, but two highly recessive sublethal mutations can also generate POD, as can be shown with equation 4.

It is not likely that a structure with lethal in repulsion emerge in a selfing species (contrary to blocks of linked weakly deleterious mutations) but it can emerge in an outcrossing species and persist after a transition from

outcrossing to selfing. So it would be interested to explore more the parameter space. At least, exploring other values of h and s would help better understanding the conditions under which POD regions are more likely to be maintained (and also to relate the work to Zhao and Charlesworth 2016). Ideally, considering both highly recessive lethal or sublethal and weakly deleterious and less recessive mutations could be useful (or even a distribution of h and s).

Another parameter which is not widely explored is recombination rate. I would guess that change in regime could depend on the ratio r/u where u is the mutation rate towards deleterious mutations.

Overall I thus recommend to reject this manuscript in the current form I encourage the authors to properly address the different issues, and especially to justify more clearly whether the proposed scenario is biologically possible or not, and under which conditions.

I have also a few minor comments:

L 118 : “This much balancing selection” : strange wording

Equation 6 : why not using more accurate expression given by Roze 2015 ?

L157 : here the term inbreeding load is used whereas it is for inbreeding depression. The term inbreeding load is usually used in a slightly different way as the slope of log-fitness with inbreeding coefficient (see Morton Crow Muller model)

L158 : “A2from” : a space is missing

Eq 8 : the use of i as a parameter is a bit disturbing at first sight because we can think about a counting index (but this is just a matter of taste)

Sylvain Glémin

Reviewed by *Yaniv Brandvain*, 17 Feb 2022

Review of: The role of pseudo-overdominance in maintaining inbreeding depression

In this manuscript, the authors explore how a “pseudo over dominant (POD) block of the genome” evolve, and how their evolution both impacts and is impacted by ongoing mutations at other regions of the genomes. Pseudo-over dominant genomic regions are portions of the genome for which heterozygotes have higher fitness than homozygotes, due to the action of numerous partially recessive mutations in repulsion phase disequilibrium, rather than the action of truly over variants. The authors consider how the answers to these questions can change depending on mating system and population size.

The work has a few take-home messages – (1) POD regions are often unstable (2) POD regions are more stable when they are made of many tightly linked mutations, and/or other regions of the genome, and/or heterozygosity is favored genome wide (3) POD regions are less stable when made of fewer more loosely linked mutations in populations are small (or effectively so because linked selection decreases N_e) and/or more selfing. (4) When POD regions are maintained high inbreeding depression can be maintained – even in largely selfing populations.

On the whole, the work seems correct, and like a nice contribution to the ongoing interest / appreciation of pseudo overdominance. However, I have numerous questions / concerns about decisions with regards to model conception and parameters choice etc. I also recommend the authors consult our recent preprint [Sianta et al. first posted may 2021, reposted Dec 2021 <https://www.biorxiv.org/content/10.1101/2021.05.20.445016v2>] which modelled the origin of POD, whose results complement this work (In hope to integrate the findings from this paper into ours).

Concerns

Beginning with a POD region

The authors start their study with a stylized POD region already existing. Their underlying model is that two isolated populations have themselves accumulated a “local drift load” composed of n equally spaced deleterious mutations per haplotype in some specific genomic region, and with no other deleterious mutations elsewhere in the genome (although they can arise after this initial starting point). The biological basis for such a model is unclear. Harkness, Brandvain, and Goldberg (2019, JEB) began with a similar premise – assuming that isolated populations each fixed their local drift load but did not assume such variants to be highly localized to a single genomic region, and it is unclear why this would occur biologically. My sense is that it would not, and that rather than being strongly biologically motivated, this model is a hack to get at the interesting questions the authors hope to explore. If this is the case, the authors should make that clear, if not, they should provide more evidence that this is a reasonable biological scenario (e.g. perhaps such regions have exceptionally low recombination rates or high mutation rates, such that the load can be purged from other genomic regions, but not these?).

Composition of POD regions

In addition to being surprised by the pre-existence of POD regions, I was confused by the properties of mutations in these regions. (as well as the ongoing mutations experienced by the population). The authors assume a dominance coefficient of 0.2 and a selective coefficient of 0.01. These choices were justified by pointing to the mean values from Agrawal and Whitlock’s efforts to estimating the distribution of these values from yeast knock outs. It is not obvious to me why this value is relevant for mutations that could make up a POD region – which should not be random draws from the 2D distribution of fitness and dominance effects – lets alone equal one value in the center of the distribution. A dominance value of $h=0.2$ is higher than what I imagine would make up PODs generated and maintained by a process other than the one considered in this model – so I worry about the applicability of these parameters.

Explanation of results

Often the explanation of the results seems pretty shallow. For example the authors state that “Counter-intuitively, POD heterozygosity persists much longer (relative to neutral heterozygosity) in smaller populations than in larger populations (compare the red and blue lines in Fig. S3 where $N = 100$ to Fig. 2 and Fig. S2 where $N = 1000$ and 5000).” I could logic why the authors found this to be counterintuitive, and I can logic an explanation for this result, but neither were provided.

Limit discussion of the work to what was found

The abstract and discussion suggest that the work is related to the formation and persistence of POD regions, however the methods and results focus only on the persistence. My suggestion is to not discuss the formation of POD regions as it is pretty far afield from what was modelled. It is also somewhat strange to have statements like “it seems unlikely that POD’s would arise in isolated populations” which depend on modelling assumptions of $h = 0.2$, etc (see above)..

Enhance clarity of figures, parameter names, etc etc

I found all the figures and parameters to be difficult to follow. Most increased my cognitive load substantially. I am afraid this could turn readers off and decrease the impact of this work. For example, denoting the population size as N , and the number of loci as n is a bit confusing, and made my life hard. Similarly, the sometimes the “control” (i.e. no POD region) was labelled as “no POD” and other times as $s = 0$. Direct labelling of figures would make results easier to process, and little changes like have the x-axis as “selfing rate (σ)” rather than “ σ ” would make it all easier to follow. Finally, I may be being dense, but I am having trouble seeing if there is any different information in the top (a & b) and bottom (c & d) parts of figure 5



Reviewed by [Lei Zhao](#), 21 Jan 2022

[Download the review](#)