Reviewer 1

I reviewed the paper entitled "The fate of recessive or overdominant mutations near mating-type loci under partial selfing", by Tezenas et al. In this interesting manuscript, the authors theoretically investigate how mutations are purged or can accumulate near a mating-type locus, with a two-locus two-alleles model. The dominance of the introduced mutation, and the mating system of the populations are the main variables affecting the purging of the mutation. The main findings are that the presence of a mating-type locus always decreased the purging probability and increased the purging time of the mutations under selfing. The authors also showed that deleterious mutations can accumulate near the mating-type locus over evolutionary time scales.

I found the introduction to be very interesting, presenting the previous models used to study similar questions, what are their limitations and what will be improved in the manuscript. It also presents the mathematical tools previously used and that will be used in the manuscript.

I liked the fact that the authors put a lot of details around the mathematical tools they used in M&M and Results sections, even if it can make the model hard to follow for non-theoretical people.

The discussion is interesting, and proposes to link theoretical outcomes with empirical data, notably in fungi. The limitation section is also nice and discusses some potential unexplored mechanisms that could modify the outcomes of the presented model.

Overall, I enjoyed reading the paper, and I only have minor comments related to the clarity of some sections of the manuscript. I hope you will find them helpful.

1. **L141: "Each individual reproduces at rate 1:" -> Consider reformulating the sentence, the beginning is confusing to me.**

   >>> We would like to keep "each individual reproduces at rate 1" because it is a classical feature of such stochastic models. We have nevertheless clarified the sentence: L147 "After a random time following an exponential law of parameter N, an individual is chosen uniformly at random to reproduce. This means in particular that all individuals have the same probability to reproduce. Mathematically speaking, this formulation is equivalent to saying that each individual reproduces at rate 1. "

2. **Section 2.1: I would explain somewhere what is the difference between intra- and inter-tetrad selfing events somewhere, as it is probably not obvious for most readers, notably regarding the production of heterozygotes offspring (it is explained in lines 386-387, but it could be described sooner).**

   >>> We have added an appendix (App. 6) to clarify the differences between intra- and inter-tetrad mating, and specified their contrasting effects on the percentage of heterozygous offspring produced in the introduction:

   L106 : "Individuals can reproduce via outcrossing, or via either one of two types of selfing, intra-tetrad mating or inter-tetrad mating. The two types of selfing depend on whether a given gamete mates with another gamete produced during the same meiosis event (within a tetrad) or with a gamete from a different meiosis (from another tetrad, Appendix 6). The distinction is important because intra-tetrad mating maintains more heterozygosity in some genomic regions than inter-tetrad mating (Hood and Antonovics, 2000)."

3. **Lines 147-149: In the case of inter-tetrad selfing, when referring to a "pairs of gametes" among all the possible viable possibilities, you only sample one gamete per tetrad right? As written it is not clear.**

   >>> We changed the description into (see L152) "(i) Intra-tetrad selfing, with probability fp_{in}: the two gametes are picked from the same tetrad, only one parent is involved; (ii) Inter-tetrad selfing, with probability $f (1-p_{in})$: the two gametes are picked from two different tetrads produced by the same individual, only one parent is involved; (iii) Outcrossing, with probability $1-f$: the two gametes are picked from tetrads produced by two different parents. In this case, the second parent is chosen uniformly at random in the remaining population, and produces haploid gametes via meiosis with the same recombination rate r."

4. **L240: Remove the "in English".**

>>> L246 We have clarified the citation into "Sewastjanov, 1975, in German, and Penisson 2010 for a statement of these results in English". We find it important to cite Penisson 2010 because Sewastjanow's book is really hard to find, and Sophie Penisson's PhD work was a key access to these results. However, it is important to mention that the second is an explanation of the first one and not a set of different original results.

5. **L280-290: By simulations, do you mean numerical iterations, or individual-based simulations? If I understood well this is numerical iterations, but I am not 100% sure. Add a little bit more detail here about the simulations performed.**

    >>> Thank you for the suggestion. We actually performed stochastic individual-based simulations. We have clarified this point on L310 : "The branching process was simulated with a Gillespie algorithm to obtain an empirical distribution for the time to extinction. More precisely, the Gillespie algorithm produces realizations of the stochastic process by iteratively updating the numbers of individuals of each genotype within the multitype branching process."

6. **L322: "the deleterious mutation is purged from the population before it reaches a significant frequency with probability one" -> "the deleterious mutation is purged from the population with probability one before it reaches a significant frequency" ?**

    >>> What we mean is that the branching process is subcritical, i.e. that the process goes extinct almost surely. However, the branching process is an approximation for the early dynamics of a mutation : if the mutation reaches a significant frequency before the process goes extinct, the approximation does not stand anymore. The probability at hand here is the probability of extinction of the branching process, which corresponds to the probability that the mutation is purged before reaching substantial frequencies. We have clarified the sentence : L351 "the probability that the deleterious mutation is purged from the population before it reaches a substantial frequency is one".

7. **L321-329: Give numerical values for the purging time. You are not referring to any figures or tables, so numerical values would help the reader to understand what "… purged … take a much longer time…" means.**

    >>> At this point, only theoretical results are used to describe the dynamics of deleterious mutations. As described in the Methods section, when rho<0, the process is subcritical and it has been shown that the probability of extinction is one, and the mean time to extinction is finite. On the other hand, when rho=0, the probability of extinction is still equal to one, but the mean time to extinction is infinite : this means that purging can take a very long time. We do not wish to give numerical values at this point, even more so as we do not give any values of the purging time for the critical case (because some simulations would take too long to run, this is also explained in the Methods section). However, we specified that the results given here relied on theory, L348-358 : "Under partial dominance, we find that the dominant eigenvalue rho of the matrix C is always negative or null (see App. 7.2.1 and 7.2.2 for more

details on the proof and computations). Previous theoretical results on branching processes state that, when rho<0, the probability that the deleterious mutation is purged from the population before it reaches a substantial frequency is one, and the mean time of purging is finite (see the Methods section). In particular, the probability of purging does not depend on the mating system (rho<0 for any value of intratetrad, intertetrad and outcrossing rates), nor on the recombination probability, selection or dominance coefficients. The only exceptions are when the deleterious mutation is neutral (s=0) or behaves as neutral (h=0 and r=0, the mutation is neutral when heterozygous and completely linked to one mating-type allele), in which case the dominant eigenvalue is 0. The mutation is still purged from the population but previous theoretical results on branching processes state that this can take a much longer time compared to the case where rho <0, as the mean purging time would be infinite (see the Methods section)."

8. **L330-348: I think this section can be shortened and moved toward the Discussion section. It is long, and even if interesting, it can be removed from here as the Results section is already long enough to avoid comparisons with previous models.**

>>> We find it interesting to keep this comparison, in order to compare the results obtained under random mating versus selfing. However, we agree with your suggestion to remove this detailed paragraph from the Result section. In order to keep the mathematical technicity level of the discussion to a minimum, we have instead created a new paragraph in the method section to describe this previous model (L257-269) and we refer to this paragraph in the Result section (L359 and 387).

9. **L367-368: Again, a numerical value would help to understand what "long" means.**

>>> Here again, following our reply to point 7, we did not give numerical values, but made a clearer reference to the Methods section where theory is explained. L374 : "The dynamics of the b-subpopulation (i.e. mutant carriers) is then critical, which means that the mutant is purged with probability 1 but the purging time can be arbitrarily long (since the average extinction time of a critical branching process is infinite, see the Methods section)".

10. **Fig.4: The mean is not very informative here, and is driven by rare, extreme times of purging. I would remove the "mean" lines to lighten the figure. In addition, you are not discussing this metric in the main text.**

>>> We removed the mean lines from histogram figures (4, S3, S4, S5), this indeed lightens the figure. A few changes were made in the text as a consequence :

L463 : "(several orders of magnitude longer than the 75% percentile empirically obtained from the 100,000 runs)" instead of "(Several orders of magnitude longer than the mean)"

L461-471 : We removed all occurrences of "Mean", and kept only "purging time"

L533 : "the purge of the deleterious mutation took several orders of magnitude longer than the mean purging time" changed into "the purge of the deleterious mutation took several orders of magnitude longer than the 75% percentile empirically obtained from the 100,000 runs"

11. **L516: "Deleterious population" -> You mean deleterious mutation? If not, the sentence is unclear.**

    >>> Yes indeed, thank you for spotting this incorrect formulation.

12. **L523: Abu Awad & Roze is not the only model investigating multi-locus dynamics of interacting deleterious mutations with different mating systems, indeed Lande and Lande & Porcher paved the way (Lande, 1977; Lande & Porcher, 2015), more recently some people added dominance to these models also (Clo & Opedal, 2021). Nevertheless, those models are a bit different because in these multi-locus models, mutation accumulation is rather positive for fitness in selfing populations due to compensatory effects.**

    >>> Thank you for these recommendations. The results on the impact of selfing presented in those papers are indeed very interesting. We have added a citation in the corresponding section of the discussion : L565 "Indeed, selfing has a non-monotonous effect depending on the tightness of linkage between multiple interacting loci (Abu Awad and Roze 2018): at low selfing rates, increasing linkage between loci increases the mutation load, whereas the opposite effect is observed at high selfing rates. Selfing also has a non-monotonous effect on genetic variation in population under stabilizing selection (Lande and Porcher 2015, Clo and Opedal 2021)."

Literature cited:

Clo, J. & Opedal, Ø.H. 2021. Genetics of quantitative traits with dominance under stabilizing and directional selection in partially selfing species. Evolution 75: 1920–1935. John Wiley & Sons, Ltd.

Lande, R. 1977. The influence of the mating system on the maintenance of genetic variability in polygenic characters. Genetics 86: 485–498. Oxford University Press.

Lande, R. & Porcher, E. 2015. Maintenance of quantitative genetic variance under partial self-fertilization, with implications for evolution of selfing. Genetics 200: 891–906. Oxford University Press.

Reviewer 2

The paper analyses how linkage to a mating type locus (which is necessarily heterozygous) influences the persistence/extinction time of a newly arisen mutation at a selected locus, focusing on scenarios where this mutation is either deleterious and (partially) recessive, or overdominant. The paper allows for arbitrary linkage between the two loci, partial selfing which may be within or between tetrads produced by an individual. The analysis focuses on the initial stochastic dynamics of the mutation and models this as a multi-type branching process (where types correspond to different two-locus genotypes carrying the newly arisen mutation). The analysis focuses on the dominant eigenvalue of the rate matrix C, which governs the long-term evolution of the expected numbers of each type, as a proxy for extinction times.

I think the paper is carefully written and well-organised. Nevertheless I have some substantive comments/queries:

1.   In general, I wonder if it might be better to show 1/rho or 1/|rho| instead of rho in the figures, as 1/|rho| has a more intuitive biological interpretation (as being rough proxies for extinction times, when rho is negative). This would perhaps also give a better visualisation of the sheltering effect: for example, rho=0.01 and 0.001 correspond to extinction times (and sheltering effects) that are different by an order of magnitude but appear as essentially the same color (e.g. in figure 2). By contrast, rho=0.1 and 0.01 (which again correspond to a difference in factor of 10) appear as very different colors. Thus, in my opinion, heat maps of rho do not give a sense of the magnitude of the sheltering effect, especially in parameter regimes where it may be most substantial.

>>> Thank you for this interesting comment. You are right when writing that when it comes to branching processes, 1/rho is usually simpler to interpret compared to rho itself. However, in our case we found that 1/rho was not a good approximation for the mean purging time (in particular because in some of the regimes we consider, the branching process is nearly critical and a given realization of the extinction time can be quite far from the mean). Moreover, rho can be interpreted as the asymptotic growth rate (or decay rate of it is negative), which is also an intuitive quantity to describe the dynamics of a population. We thus chose to keep studying rho and its derivatives.

Regarding the comment on the lack of difference in colors for values of rho with several orders of magnitude of difference, we tested a log-scale for Fig 2 (see below), in order to see whether the values of rho could be better represented (especially when changing orders of magnitude). It appears that, rather than allowing a better discrimination between the values, a log-scale gathers more values under the same color. We therefore kept the linear scale,
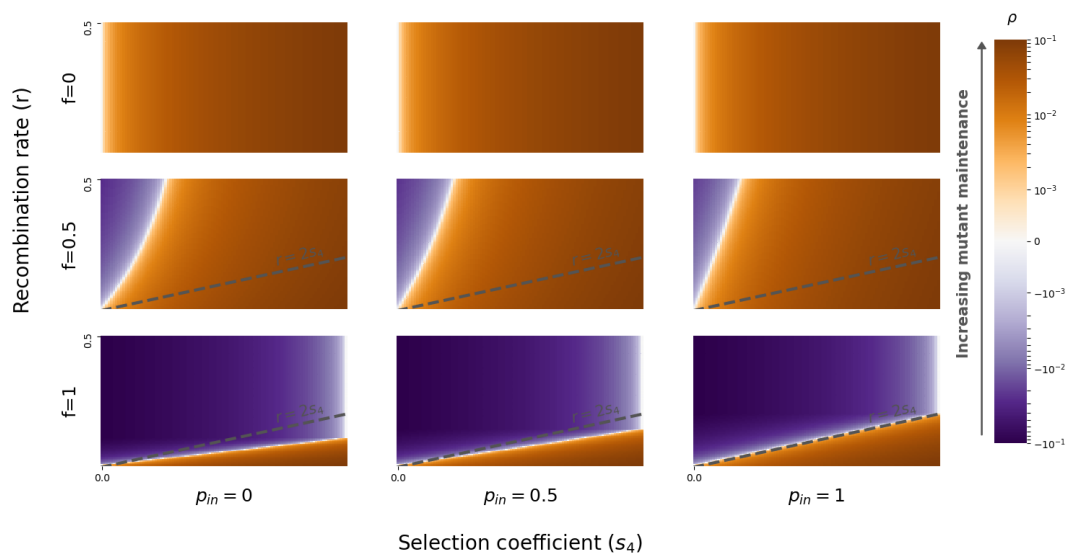
2. **Lines 311-316: With a recombination rate r0 per base pair and d base pairs, the assumption that the total recombination rate r=r0*d breaks down beyond r~0.1-0.2. By the same logic, it might be better to have recombination rate rather than base pairs along the x-axis of fig 5.**

   >>> If we understood your comment correctly, you meant that, when the distance between two loci is large, the probability that there is only one recombination event decreases, and therefore that the "effective" recombination rate is not linear any more. We changed the x-axis of Figure 5, which now represents the rate at which recombination occurs on the portion of the genome lying between the 2 loci (instead of the distance in base pairs between the 2 loci). We have corrected the text accordingly (Methods section L340-343 and Results section L472-482).

3. **If the goal is to understand whether the sheltering effect is stronger in the vicinity of a mating type locus, would it not be more meaningful to look at the derivative of rho with respect to r as r→ 0, rather than r→ 0.5. Evaluating the derivative at r=0.5 essentially corresponds to comparing the effect of two locations that are both quite far from the mating type locus. For example, a chromosome of 100 cM (respectively 150 cM) corresponds to a recombination fraction r~0.432 (resp. r~0.475) between the two ends of the chromosome (assuming Haldane's map function). Even if one allows for crossover interference, r=0.5 would still correspond to rather large map scales, that are not so relevant to questions considered here (about the formation of strata and sheltering effects in the vicinity of the mating type locus)?**

>>> Our aim in this paragraph is to study the impact of the existence of (even weak) linkage to a mating-type locus on deleterious mutation persistence, and therefore we found it relevant to study the derivative at r=0.5. However, you are right in saying that the derivative at r=0 gives information on the impact of the linkage strength on mutation persistence. To make the difference between the 2 questions clear, the derivative at r=0.5 gives the impact of introducing a slight amount of linkage to a mating-type locus on the dynamics of the deleterious mutation, while the derivative at r=0 gives the impact of departing from complete linkage to a mating-type locus (by introducing a small amount of recombination) on the dynamics of the mutation.

The computation of the derivative at r=0 gives us that this quantity is always negative, in both selection scenarii. We have added the calculations in Appendix 8.2.3 L999 and 8.3.3 L1063.

We have also plotted the difference between the two derivatives (at r=0.5 and r=0) as a heatmap that is displayed in the supplementaries (Fig S3). We introduce this comparison in the Methods section (L292 and below), and discuss the results in the Results section (L448).

Methods L292 : "We also look at the strength of the sheltering effect on mutations close to the mating-type locus, by studying the eigenvalue variation around r=0. Setting the recombination rate to r=0 models a situation where the load locus is completely linked to the mating-type locus. Hence, the mutation is completely linked to one mating-type allele, and maintained heterozygous. Looking at the derivative d rho/dr |r=0 allows us to quantify the impact of departing from this situation by loosening the linkage between the two loci. We study the difference between the derivative at r=0.5 and the derivative at r=0 to compare the effect of adding a small amount of linkage between completely unlinked loci (r=0.5) and the effect of adding a small amount of recombination between completely linked loci (r=0)."

Results L448 : "Looking at the derivative at r=0, we show in App. 8.2.2. and App. 8.3.3. that the derivative at r=0 is also negative in both selection scenarii. This means that the eigenvalue decreases, i.e. that the mutation is less maintained in the population as soon as the two loci are no longer completely linked. Figure S3 shows that the difference Delta rho is always positive, which means that the absolute value of derivative at r=0 is larger than the absolute value of the derivative at r=0.5. This shows that the sheltering effect of the mating-type locus is stronger on closely located mutations : adding a small chance of recombination on previously completely linked loci (r=0) has a greater impact on the maintenance of deleterious mutations than adding a small amount of linkage between two previously completely unliked loci (r=0.5). The largest difference between the two derivatives occurs for selfing rates close to one. For those rates, the derivative at r=0.5 is zero, whereas the derivative at r=0 approaches -1. This shows that the linkage to the mating-type locus particularly impacts the strength of its sheltering effect under high selfing."

4. **Line 341: Based on the notation above, w23 should perhaps be w32; same comment for line 369.**

   >>> Yes indeed, thank you for spotting this.

5. **Overdominance scenario (line 349 onwards): It may be interesting to discuss the effect of the mating type locus on unlinked (r=0.5) loci from the point of view of it being a baseline "genomewide" effect.**

    >>> The genome-wide impact of a mating-type locus (or a permanently heterozygous region) is indeed an interesting question, but our model has not been designed to answer it. Indeed, we would have to compare our results to results for populations in which there is no mating-type at all, as having a mating-type constrains the formation of offspring and thus the segregation of other loci during meiosis/mating. A branching process model can still be used, but all the rate calculations (tables shown in Appendix 7) would have to be computed, and this would be a completely different study.

    We have added a comment in the part of the discussion on perspectives to highlight this point : L599 "The diversity of observed patterns regarding the presence or absence, length and number of evolutionary strata around these regions (Uyenoyama 2005) may be explained, in addition to the mating system, by other factors controlling the long-term behavior of deleterious mutations which are not studied here, such as the number of alleles at supergenes, the length of the haploid phase (Jay et al 2022), or the presence of multiple load loci that are possibly physically linked and with epistatic interactions (Abu Awad Roze 2018, Lenormand Roze 2022). The questions of the genome-wide impact of a mating-type locus, and of the interaction between a permanently heterozygous locus and background mutations, are currently debated (Abu Awad Waller 2021). The branching process framework developed here could be applied to diploid individuals carrying a load locus with two alleles, undergoing selfing or outcrossing, in order to investigate the dynamics of a new deleterious mutation in a population with or without a mating-type locus."

6. **I was also curious whether one can come up with a clean biological interpretation for rho to be negative in this r=1/2 limit (for 0<f<1): based on App 7.3.2. Based on Fig 2, rho appears to switch sign at some threshold value of s4: can we understand this threshold in a more intuitive way for the r=1/2 case?**

    >>> We looked at the explicit value of rho for r=0.5, but there are no neat simplifications unfortunately... It thus seems complicated to obtain a clean and general biological interpretation. Figure 2 however shows that, for high values of s4, rho > 0 : when the homozygous wild-type genotype is highly disfavored compared to the heterozygotes, the overdominant mutation is maintained longer in the population.

7. **Lines 359-362: This does not really correspond to a special case of over-dominance, but just describes a completely recessive b allele.**

  >>> We agree: rho has the same value in the Partial Dominance case when h=0. We have added: L379 "This corresponds to a completely recessive mutation, and is in agreement with the results for the partial dominance case with h=0." We still find it interesting to explicitly state this particular case, because it has a particular meaning in the overdominance case : we suppress selection against the wild-type genotype, whereas in the partial dominance scenario, we suppress selection against the heterozygotes.

8. **Figure 2: In the f=1 panel (bottom row), there appears to be a discontinuous change in the blue shades at a critical value of r, independent of s4: (e.g., around r=0.1 for pin=0, around r=0.2 for pin=1 etc.). Is this an artefact of the color scale or a real effect- and if it is the latter, then what does it correspond to?**

>>> This is a real effect. In the overdominance case, when f=1, if r.alpha - 2s3 >0 (a condition that depends only on r, s3 and pin), then rho = s4-s3. It corresponds to the upper part of the bottom panels, the color varies only with s4 (s3=0,1 in this figure). The threshold is a bit tricky to interpret (as above, formulae are not easy to parse…), but Figure 2 gives some understanding : if the recombination rate is larger than the selection strength against deleterious homozygotes, then the dynamics of the mutation relies only on selection coefficients. The threshold on recombination becomes higher as pin increases : the linkage of the mutation to the mating-type locus has the highest impact under intra-tetrad selfing. This confirms the main result of the paper, i.e. that the sheltering effect of the mating-type locus is greatest under intra-tetrad selfing.

We have added these comments in the main text and in the appendices :

L396 : "In the case of complete selfing (f=1), we find that rho = s_4-s_3 <= 0$ when r(2-r-p_{in}(1-r))-2s_3 >= 0. This shows that the dominant eigenvalue depends only on the selection coefficients when the recombination rate r exceeds a certain threshold (visible on the bottom panels of Figure 2). This means that, if the recombination rate is larger than the strength of the selection against deleterious homozygotes, the mutation is purged with probability one. Moreover, the purging time is shorter when the difference in fitness between the two homozygotes is larger. The threshold on recombination increases as p_{in} increases, which means that the strength of the linkage between the mating-type locus and the mutation has the highest effect under intra-tetrad selfing."

L1030 and onward : "Under complete selfing (f=1), if r(2-r-p_{in}(1-r))-2s_3 >= 0, then rho = s_4-s_3 <= 0. This shows that the value of the dominant eigenvalue, and thus the dynamics of the process, depends only on the selection strength when the recombination rate r exceeds a certain threshold. Moreover, this threshold depends only on the selection

coefficient for homozygous deleterious (s_3), and on the probability of intra-tetrad mating (p_{in}). This threshold appears on the bottom panels of Figure 2."

9. **Section 3.2: I think in general it may be good to distinguish more carefully between how the presence of a mating type locus affects sheltering vs. how (tight) linkage to a mating type locus affects sheltering. The first question is related to how the presence of a mating type locus changes dynamics at an unlinked (r=0.5) locus in comparison to a situation where there is no mating-type locus. For instance, how different is the largest eigenvalue (as calculated here) or extinction time of a deleterious allele from that in a single-locus model with only the load locus (and no mating-type locus).**

>>> Thank you for this comment. We have performed the following changes :

In the Methods section, we have clarified the interest in looking at the derivative of the dominant eigenvalue at r=0.5 in the paragraph starting at L275. We changed the sentence "Moreover, in our model, setting the recombination rate to $r=0.5$ allows to consider a load locus completely unlinked to the mating-type locus, while decreasing the value of r increases the strength of linkage between the two loci." into "Moreover, in our model, setting the recombination rate to $r=0.5$ allows us to consider a load locus completely unlinked to the mating-type locus, while decreasing the value of r introduces some loose linkage between the two loci."

L284 "i.e. when linkage between the two loci increase" changed into "when linkage between the two loci appears"

We have also added a new paragraph to this section in order to introduce our study of the derivative at r=0 :

L292 : "We also look at the strength of the sheltering effect on mutations close to the mating-type locus, by studying the eigenvalue variation around r=0. Setting the recombination rate to r=0 models a situation where the load locus is completely linked to the mating-type locus. Hence, the mutation is completely linked to one mating-type allele, and maintained heterozygous. Looking at the derivative d rho/dr |r=0 allows us to quantify the impact of departing from this situation by loosening the linkage between the two loci. We study the difference between the derivative at r=0.5 and the derivative at r=0 to compare the effect of adding a small amount of linkage between completely unlinked loci (r=0.5) and the effect of adding a small amount of recombination between completely linked loci (r=0)."

The paragraph in the Results section describing the study of the derivative at r=0 highlights once more the difference between the existence of a sheltering effect, and the strength of this effect for tightly linked mutations :

L448 : "Looking at the derivative at r=0, we show in App. 8.2.2. and App. 8.3.3. that the derivative at r=0 is also negative in both selection scenarii. This means that the eigenvalue decreases, i.e. that the mutation is less maintained in the population as soon as the two loci

are no longer completely linked. Figure S3 shows that the difference Delta rho is always positive, which means that the absolute value of derivative at r=0 is larger than the absolute value of the derivative at r=0.5. This shows that the sheltering effect of the mating-type locus is stronger on closely located mutations : adding a small chance of recombination on previously completely linked loci (r=0) has a greater impact on the maintenance of deleterious mutations than adding a small amount of linkage between two previously completely unliked loci (r=0.5). The largest difference between the two derivatives occurs for selfing rates close to one. For those rates, the derivative at r=0.5 is zero, whereas the derivative at r=0 approaches -1. This shows that the linkage to the mating-type locus particularly impacts the strength of its sheltering effect under high selfing."

10. **Figure 3: It wasn't clear to me what is meant by the values of the derivative- "scaled by their minimum"- what does the minimum refer to here? I was also wondering whether it'd be more natural to show the derivative (at r=0.5) scaled by the value of rho at r=0.5, since this gives an indication of the relative change in rho. However, more importantly: see my comment above on evaluating the derivative of rho w.r.t. r at r=0.5 vs. r=0.**

>>> The suggestion to rescale by the value of rho was a good one. However, a problem is encountered when rho = 0 (which occurs in the overdominant scenario). The heatmap for the partial dominance scenario is very similar regardless of whether we rescale the values of the derivatives by the value of rho or by the minimal value the derivatives take over the parameter range explored here (see below). We therefore chose not to rescale by the value of rho, to avoid complications in the overdominance scenario.
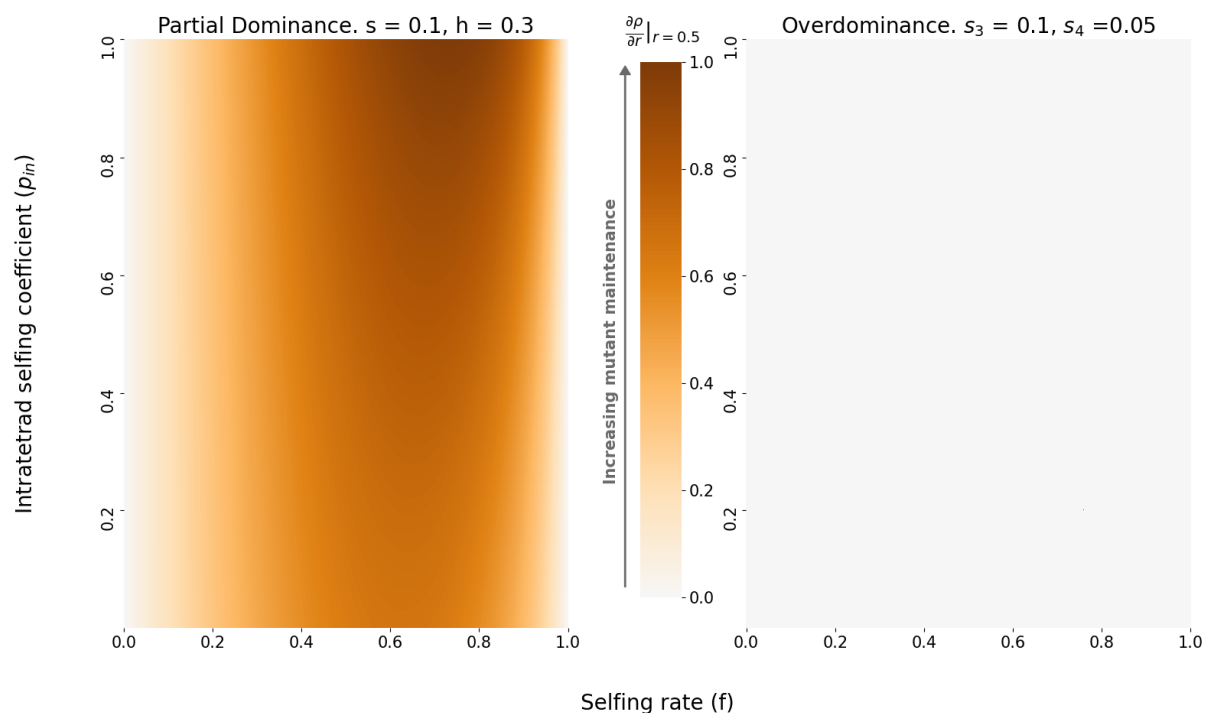
The figure shows the derivative scaled by the value of rho at 0.5 for the same parameters. For the partial dominance scenario (left), the shades are very similar. For the overdominant scenario (right), when rho is nearly equal to zero, rescaling flattens all other values, which gives this nearly-white heatmap.

However, we clarified what we meant by "scaled by their minimum" in the caption of Figure 3, and changed it into : " For each panel, the values of the derivative range from a minimal value, which is negative, to zero. We divided each value of the derivative by this minimum in order to plot values between 0 and 1 for every panel. This enables us to compare the effect of the presence of a mating-type locus on the same scale for both selection scenarii."

We have performed similar changes in the captions of Figures S1, S2.

11. **Figure 4: Perhaps it may be more insightful to show this as a semilog plot..**

>>> The x-axis was already on a log scale, we have clarified this point by adding "The x-axis is log-scaled" in the captions of Fig 4, S4, S5 and S6.

I understand that the distribution is not predicted by the Gumbel law, but what sort of distribution is this? Is it a power law (at least in some regime of t), for example, as one would expect for a critical BP?

>>> We have tried to fit the empirical distribution by some forms of densities. Apparently, a good approximation for the left part of the histogram is given by a power law (see the figure below, in which the density is $1/(4+x^{1.7})$). However, the dominant eigenvalue does not appear in an obvious way in this expression and the role it plays in the shape of the density is hard to identify.
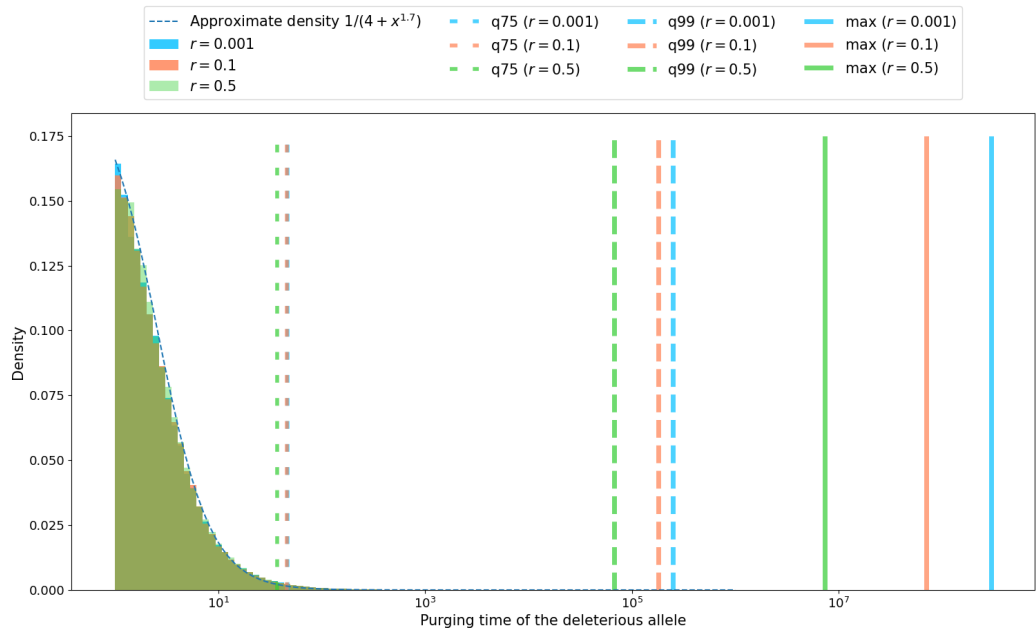
Also what is the value of rho in this figure: i.e., how close is it to zero?

>>>This is indeed a good question. We have added the values of rho for each histogram (Figs 4, S4, S5, S6) in their captions, and adapted the main text accordingly :

L465-467 : Note that the approximation of the distribution of the time to extinction by a Gumbel law (Th. 4.1 of Heinzmann 2009) falls short here, because the initial number of individuals (one) and the absolute value of rho (given in the caption) are too small.

>> In each case, the closer rho is to zero, the more extreme the rare events are : the distribution of the 1% longest purging times is stretched towards higher values when rho becomes closer to zero, while the distribution of the 75% shortest purging times does not change much.

Reviewer 3

This preprint outlines branching-process based results to demonstrate how recessive and overdominant deleterious alleles are sheltered under mating-type loci under different types and levels of self-fertilisation. Results are presented using both analytical formulae and simulations, which uncover novel behaviour regarding how deleterious mutations could be sheltered for long periods of time. For example, the model is used to investigate the trade-off between the purging of deleterious mutations under selfing, compared to the maintenance of heterozygosity due to linkage to the mating-type locus.

This is a great preprint that clearly outlines the model derivations, along with the key results while using numerical computations to illustrate important outcomes of how recombination and selfing interact to maintain deleterious variation around mating-type loci. It is also clearly written, with the theoretical assumptions neatly laid out and derivations well-explained – I found it a genuine pleasure to read.

1. My only major concern is that I found some of the biological justification of the model to be lacking. It looks at deleterious mutation dynamics around mating-type alleles under different levels of selfing, including complete selfing. However, I have trouble understanding how there can be a combination of complete self-fertilisation and mating type loci in a population, since the presence of mating-type loci should restrict selfing by definition? Perhaps the authors could provide some further biological justification as to how the two cases could co-exist, or whether these results hold under other types of balancing-selection type cases under selfing.

>>> We have added an appendix that describes the three mating systems, and we have (hopefully) better explained all along the main text that mating types as modeled here (and as found in fungi for example) do not prevent diploid selfing :

L74, already present in the previous version : "Indeed, because mating types are determined at the haploid stage in fungi, mating types do not prevent selfing when considering diploid individuals (Billiard et al., 2012)."

L133: "on the long-term maintenance of deleterious mutations near a fungal-like mating-type locus (i.e., not preventing diploid selfing). "

L13: "presence of a fungal-like mating-type locus (i.e., not preventing diploid selfing) always sheltered the mutation under selfing"

L140: "We model a fungal-like mating-type locus, so that mating is only possible between haploid cells carrying different alleles at the mating-type locus (this does not prevent diploid selfing as each diploid individual is heterozygous at the mating-type locus)".

L487: "We have shown that partially recessive deleterious mutations close to a fungal-like mating-type locus (i.e., not preventing diploid selfing)"

L593 : "The results obtained here on the accumulation of deleterious mutations should apply, beyond fungal-like mating-type loci, to other permanently heterozygous loci, such as supergenes (Llaurens et al. 2017). In contrast, sporophytic plant self-incompatibility loci prevents diploid selfing, leading to a completely different evolutionary scenario in their flanking regions as imposed by complete outcrossing"

Minor comments (L = line number):

1. L2: change "having extended with time" to "having extended over time".

>>> Done

2. **L199: Explain what N^3 is for non-mathematical biology readers.**

    >>> We have added "(vectors with three integer-valued coordinates)" after N^3.

3. **L206: change 'descendance' to 'descendant'**

    >>> We performed the following change :

    L213 : "the replacement of the parent by its descendants". Here we need to specify that there can be several descendants.

    However, we kept the terminology "descendance vector" afterwards, in order to clearly specify that this vector contains the whole descendance of an individual, that can be composed of multiple descendants.

4. **Figure 4 : it's hard to see the different distributions of time for each recombination rate as they overlap. Perhaps they could be made transparent in some way so it's clearer to see how they compare?**

    >>> Thank you for this suggestion. We have changed the transparency of the colors. It is indeed easier to see that the distributions below q75 are very similar for all sets of parameters.

5. **Appendix 6, Tables 2-4: In these table entries, I propose removing the 'X' to denote multiplication. For in-text formulae it is not too intrusive but given the amount of algebra on display in these tables, then including such multiplication symbols makes each cell harder to parse. On a positive note, I found the inclusion of these tables a great addition to the manuscript, and they helped aid me in understanding the model derivations.**

    >> Thank you for this suggestion. We removed the "x" sign in the tables to improve the readability.

6. **Page 34, set of equations just under L828: on the 5th line of equations, I think there's a typo: the last entry is meant to end with s^u and not s^v.**

    >>> Indeed, thank you for spotting this.