

Associate Editor

Thank you for submitting your article to *Evolution*. It has now received two reviews, both of which recognize its valuable contributions and essentially recommend publication after relatively minor revision. I share these views. While both reviews are positive, they make a number of worthwhile suggestions for you to consider for revising the manuscript. One reviewer has suggested combining the Results and the Discussion into a single section to avoid unnecessary discussion of the results in the current Results section. Personally, I would favor retaining both sections, **moving more discursive text from the Results to the Discussion**. I leave the decision to you, but I do agree with the reviewers that some resorting of the Results text is needed. I also agree that the **Introduction could be better motivated by referring more to empirical work**, and the **Discussion section could relate the results more to other theoretical findings** in the literature. Overall, think that revision will be relatively straightforward, but I would ask that you consider all the suggestions made by the reviewers and, if you disagree with them, that you justify and explain your reasons in a response letter. (emphasis mine)

Thanks a lot for your positive appraisal.

First of all, I wish to note that there was an important error in my original submission, which I outline below. Evidently, I have redone all simulations and revised the manuscript accordingly. The error affects both diploid and tetraploid establishment similarly, so does not affect the results qualitatively (since these are concerned, mostly, with relative establishment probabilities), but does affect all results with nonzero selection intensity ($\gamma > 0$). I outline the details of the error in the next section.

I have now revised the Introduction to refer more to empirical work. I have included the references suggested by Reviewer 1. Following the suggestions of Reviewer 2, I have tried to refer better to empirical work on (1) polyploid establishment at range margins and (2) the genetics of adaptation to peripheral environments (in polyploids). In the interest of brevity, I do not present an in-depth review of these topics. For (1), I refer mostly to Griswold (2021), who presented a nice overview. For (2), I mainly tried to better motivate my interest in *polygenic* adaptation, referring to recent work and reviews that suggest that local adaptation is typically polygenic, and work that shows that autopolyploids present no exception to this.

Regarding the Results and the Discussion sections: I do not want to do away with Discussion section altogether (as suggested by Reviewer 1), as I wish to discuss factors relevant to the problem which are not studied in my article (dominance, inbreeding depression, *etc.*) and hence do not feature in the results section. Furthermore, I want to relate my work to other recent theoretical work (in particular that of Griswold (2021)). I believe this warrants a dedicated Discussion section. I have sought to move ‘more discursive text’ to the Discussion section, although I found there was fairly limited scope to do so when I seek to keep the interpretation of the results more or less self-contained.

I have expanded the discussion slightly to address some of the remarks of Reviewer 2 concerning the issues of mutation and the scaling of genetic variance. I have added a concluding paragraph following the suggestion of Reviewer 1 to highlight that other eco-evolutionary questions about mixed-ploidy populations might be fruitfully addressed using the infinitesimal framework. I took the opportunity to allude to the work of Oswald and Nuismer (2011) in that section, following up on a suggestion of Reviewer 2.

Overall, I think these revisions have contributed considerably to a better-motivated article that engages more with relevant empirical work. I hence thank the reviewers and the editors for their suggestions. Below I include a detailed reply to both reviewers.

Error in the first submission

In the simulations included in the first submission of my article, I did not sample the trait values of the next generation correctly when there was nonzero selection $\gamma > 0$ (as in figs. 2-5). As outlined in my paper, I sample the number of offspring of a particular combination of gametes from a particular parental pair after selection in proportion with the expected fitness using eq. 11 (number in new version). In the previous version, I then sampled, incorrectly, offspring trait values from the Gaussian distribution of offspring implied

by the mixed-ploidy infinitesimal model (eq. 7). This however fails to account for selection on the trait value within each parental pair.

The density of trait values after selection is found easily. Writing $f(z)$ for the Gaussian density of trait values with mean \bar{z} and variance V , and $w(z)$ for the fitness function, the density after selection $f^*(z)$ is

$$\begin{aligned} f^*(z) &= \frac{w(z)f(z)}{\int w(z)f(z)dz} \\ &\propto \exp\left(\gamma(z - \theta) - \frac{(z - \bar{z})^2}{2V}\right) \\ &\propto \exp\left(-\frac{1}{2V}(z^2 - 2z\bar{z} - 2V\gamma z)\right) \\ &= \exp\left(-\frac{1}{2V}[z^2 - 2z(\bar{z} + \gamma V) + (\bar{z} + \gamma V)^2 - (\bar{z} + \gamma V)^2]\right) \\ &\propto \exp\left(-\frac{1}{2V}[z - (\bar{z} + \gamma V)]^2\right) \end{aligned}$$

i.e. $f^*(z)$ is a Gaussian density with mean $\bar{z} + \gamma V$ and variance V . The correct trait value distribution for the offspring in parental pair (i, j) with gametes of ploidy k and l *after selection* is hence a Gaussian with mean $\bar{z}_{ij}^{kl} + \gamma V_{ij}^{kl}$ and variance V_{ij}^{kl} .

The incorrect results from the previous submission can be seen as a kind of zeroth order approximation to the correct results: i.e. they implement selection on the mean trait value across families, but not selection on the variance. The main consequence of this error is that adaptation is more efficient than in my previous results, and hence establishment is easier. This is however the case for both diploids and tetraploids. The effect on establishment is however nonlinear, so although the main qualitative results of the paper remain the same, the new results do not simply amount to a rescaling of the old ones.

I apologize for this oversight.

Reviewer 1

This interesting manuscript investigated how the infinitesimal model can be expanded to polyploidy populations (triploids and tetraploids here). It proposes a mathematical formulation of a polyploid population's expected genetic diversity for a given amount of inbreeding. The model is then used to study how polyploidy can affect the adaptation to a new habitat (based on a sink-source model). I found the manuscript to be interesting and only have a few things to say about it. I will for sure motivate further expansion of the model to understand the evolution of autopolyploidy. Having no line number does not help to comment precisely on the manuscript; I did my best to be as precise as possible.

Thank you for these encouraging words, and my apologies for the oversight when it comes to line numbers!

Results are already well discussed when presented. I would either merge the results and discussion sections (as in brief communications for example), because it reads well as presented, or remove any explanations/discussion from the results to put them in the discussion section. The first option sounds better to me.

I appreciate the comment and have tried to rework the text with this in mind. However, following the suggestion of the associate editor, I went with the second option. I refer to my answer to the associated editor above.

A small concluding section where the authors mentioned other questions to study with their model will be useful in my opinion. The author proposed to complexify the current model of adaptation with dominance or to relax some hypotheses, but I think other biological scenarios could be studied too.

Whereas I don't think listing potential biological scenarios that can be investigated using my modeling approach is very useful, I think the reviewer is right in that it may be useful to highlight that the modeling approach can be employed to address other questions related to the ecology and evolution of mixed-ploidy populations. I have now added a small paragraph at the end that does so.

Defining establishment when $N = 100$: Is it a classic way to define establishment in such models? Or does it come from your experience with the model and you see that Under $N=100$ extinction remains likely? More details would be needed here.

$N = 100$ is indeed a fairly arbitrary choice. This threshold was also adopted in Barton and Etheridge (2018), so it makes comparisons of establishment probabilities *etc.* with their paper somewhat more straightforward. $N = 100$ seems to be a reasonable threshold in that extinction becomes sufficiently unlikely, while we do not have to simulate large populations (our approach requires tracking the $N \times N$ matrix of identity coefficients). Also, for the migration rates considered, $N = 100$ is sufficiently large that migration alone (without population growth due to adaptation) can not lead to $N = 100$, i.e. the population does have to increase the mean trait value in order to reach establishment. I did a check by putting the threshold at $N = 150$, and the establishment probabilities do not change appreciably.

Abstract: I am not sure the reference to Barton et al. (2017) is appropriate here. Maybe a more general definition would fit better.

The reason for the reference is that there appears to be some confusion about different models that bear the name 'infinitesimal model' (as outlined in for instance Turelli (2017) and Walsh and Lynch (2018)). I wanted to make it clear that we use the "Gaussian descendants" infinitesimal model, which is very carefully outlined in Barton, Etheridge, and Véber (2017). However, I guess this is sufficiently clear from the introduction and the methods sections, so the reference in the abstract can indeed be omitted, as it is in the revised manuscript.

Introduction: P2 (lower inbreeding depression (Ronfort 1999, Otto & Whitton 2000): Some empirical data? Husband et al. (2008), Clo & Kolar (2022)).

I have now added the suggested references, thank you.

Model and methods: P3-P4 (mixed ploidy population model): maybe write that the rate of reduced gametes is fixed and genetically inherited from diploids to tetraploids. It is clear from the equations but I suspect it could be clearer for non-theoretical people who will not necessarily focus that much on equations or matrices.

I added a sentence below eq. 1 to emphasize this. I note that this is also emphasized in the discussion "In addition, we have assumed a relatively high and constant rate of unreduced gamete formation u and triploid fertility v in all our simulations (5%), whereas these are known to be variable across the population, and at least in part genetically determined (Kreiner et al., 2017a; Clo et al., 2022)"

P4 (tetraploids are not twice as big as diploids): Some references would be needed to justify (Porturas et al. 2019 for example).

I have added the reference to Porturas et al. (2019).

Discussion about inbreeding depression: again, some empirical data are available to support some statements (Husband et al. (2008), Clo & Kolar (2022)).

I have added the suggested references.

Reviewer 2

This paper is an important foundational work in the area of theoretical polyploid evolutionary genetics. It presents an infinitesimal model of trait evolution for triploids and tetraploids, and then applies the model to assess conditions under which a tetraploid population will establish in a peripheral region connected to a central core by dispersal. In terms of the infinitesimal model, several aspects are very helpful

- Demonstrating how genetic effects are scaled across triploids and tetraploids such that segregating variance is equivalent across ploidies, including with diploids
- How to combine the genetic effects in an individual with mixed-ploidy ancestry
- Calculating inbreeding and co-ancestry values for mixed-ploidy individuals
- Calculating the effect of unreduced gamete formation on segregation variance across diploids, triploids and tetraploids

Other informative analyses include recursive or Markov models of null expectations for frequencies of ploidies, times to diploid ancestry and effective population size across a population of mixed-ploidy. Besides this set of results, the paper then examines conditions under which a tetraploid population is established in a peripheral population connected to a central core population that is predominantly composed of diploids. Given the additive nature of gene-action in the infinitesimal model, this analysis will likely be of interest to empiricists because additive effects are thought to be the least constrained basis for something to evolve and there is empirical evidence that the genetic basis of polyploid establishment in some species is additive. This is in contrast to the theoretical results of Griswold (2021), which highlighted the potential for a recessive-basis for autotetraploid establishment in a peripheral region. Main comments focus on mutation and adding background/context.

I thank the reviewer for this appraisal.

Mutation: For low migration rates, the possibility that autotetraploid relative to diploid establishment is driven by differences in how inbreeding accumulates is interesting (figure 4), but the model does not include mutation. New mutation dissociates identity by descent from identity in state and likely slows the rate of loss of segregation variance. It would be helpful to address how scaled mutation may affect dynamics of segregating variance between diploids and autotetraploids and whether mutation may affect results. Is there a relationship between μV_m (Barton et al. 2017) and m and V on establishment? Recognizing the point seems to be important, but its analysis can be for later.

I have not considered mutation at all so far because I assumed it would not be relevant for the population sizes and timescales considered in my study. Specifically, any individual at the time of establishment derives from a completely outbred migrant individual a relatively short time in the past (10 to 100 generations, say), so the contribution of new mutation to differences in establishment probability between diploids and tetraploids should be negligible. In more quantitative terms: we assume $1/\mu \gg \tilde{T}$ where \tilde{T} is the time between establishment and the arrival of the migrant from which the established population derived. In the presence of migration the variance contributed by mutation is even more certainly negligible. As long as mV is sufficiently larger than μV_m , we will have that, for every mutation contributing a very slight decrease in the loss of segregation variance on the island, there will be many more migrant arrivals, which each introduce a completely unrelated genotype. In all our simulations with migration we assume mV considerably larger than conceivable values of μV_m .

Of course, when considering the mixed-ploidy infinitesimal model independently of the establishment question, mutation does become a relevant topic. I mention this in the revised discussion but I think it is out of scope to discuss this in any detail in the paper. I think that my approach for scaling across ploidy levels that is explicitly based on underlying allelic effects in a finite-locus model extends readily to mutation. The main difference between ploidy levels is of course that the mutational target size is proportional to the ploidy level, so that, unless the mutational variance is exactly halved in tetraploids, the contribution of mutation to genetic variance (μV_m) in tetraploids is larger than in diploids.

I have added a note in the discussion regarding mutation and why I have ignored it.

In the intro providing empirical examples of polyploid establishment in peripheral habitats would broaden readership and connections to the literature. Also, reviewing empirical work on the genetics of adaptation to peripheral environments in polyploids would broaden connections to the literature.

I agree with the relevance of empirical work on polyploid establishment and adaptation to marginal habitats

for the introduction of the paper. However, I think that to present an actual review of empirical studies here is a bit out of scope. Griswold presents a brief review in his 2021 paper, and I prefer to refer to his paper instead of paraphrasing his paragraph. Specifically, I now write: “Many empirical studies of mixed-ploidy populations find that polyploids established in peripheral habitats at the edge of a species’ range (reviewed in Griswold (2021)), and this is in accord with large scale biogeographical patterns (Rice et al., 2019)”.

While I think it is out of scope to review in any depth the empirical literature on adaptation and polyploidy, I agree that the paper lacked a clear connection to such work, and in particular I noticed that it lacked a clear motivation for the interest in *polygenic* adaptation. I have now tried to better motivate the likely polygenic nature of local adaptation and have added a couple of references to recent work on the genetics of adaptation in autopolyploids: “More often than not, local adaptation is polygenic in nature (Pritchard and Di Rienzo, 2010; Barghi et al., 2020; Bomblies and Peichel, 2022), involving many weakly selected variants across the genome, and adaptation during polyploid establishment is unlikely to be an exception. Recent studies on local adaptation in autopolyploids indeed tend to find a polygenic basis of adaptation (Bohutínská et al., 2021; Konečná et al., 2021, 2022), however it is not clear how observed adaptive differentiation in established tetraploid populations relates to adaptation that occurred during initial establishment”.

The discussion could relate findings to other theoretical works more. For example, it does not go back and place the results in the context of Oswald and Nuismer (2011), which assumed a two-locus model and varied dosage effects, as well as competition between diploids and tetraploids, including no competition. In addition it allowed for assortative mating.

I agree that Oswald and Nuismer (2011) could be a relevant paper to relate my work to, however it should be noted that this paper focuses on establishment in sympatry, and does not deal with migration, nor establishment in a new environment (they do study establishment after environmental change, but as far as I can tell they assumed a constant population size, so this is quite different from establishment in a new habitat). While my study is clearly related to theirs in that it is concerned with establishment of autotetraploids in mixed-ploidy populations, it is not straightforward to explicitly relate the results beyond noting that (1) they investigated more processes (assortative mating by trait value, competition, stabilizing selection and inbreeding depression) and (2) also found that assortative mating and selfing may promote establishment. The latter was of course already found by many authors before Oswald and Nuismer (2011)...

However, I have now added a more explicit reference to Oswald and Nuismer (2011) in the concluding note (see the associated remark of Reviewer 1), to highlight that it would be possible to do a study that is comparable to theirs but assuming infinitesimal quantitative genetics instead of a model with a small number of loci.

It would be informative to expand on the the consequences of different scalings of genetic effects. This paper made the sensible choice to scale such that segregation variance is equal between ploidies as a baseline. Griswold (2021) scaled genetic effects such that autotetraploid and diploid individuals have equal fitness as a function of proportional allele count in an individual. A consequence of Griswold’s scaling is autotetraploids have lower additive genetic variance at HWE (Gallais 2003, p. 185), so are expected to respond more slowly in the absence of another process. Gallais (2003, p. 186) mentions some studies that compared additive genetic variance between diploids and autotetraploids.

The consequences of different scalings is taken up in the results and discussion throughout the paper where I believe this is relevant. In the methods section the relationship between allelic effects and additive genetic variance is stated explicitly. In the results shown in Fig. 2, the effect of scaling assumptions on establishment probability is explicitly shown and described. In the section on establishment with recurrent migration, the phenomenon that polyploid offspring is more extreme on average when allelic effects are not exactly scaled by one half (as in Griswold) is taken up. I now added some further discussion, referring explicitly to Griswold’s scaling in the Discussion section. I now also refer to Porturas et al. (2019) and Gallais (2003).

Neither the intro nor the discussion revisit what seems to be an important point in Barton & Etheridge (2018, p. 111) about local adaptation via infinitely small effects.

I assume the reviewer is referring to the following issue: classical single locus population genetics predicts that

one needs $m < s$ for selection to maintain a locally beneficial allele when migrants introduce the deleterious allele. So, when adaptation is due to alleles of very small effect, this seems to imply that there is little scope for local adaptation, since locally beneficial variants will be swamped by gene flow.

There are multiple reasons why this is not the case in nature. The argument outlined in Barton and Etheridge (2018) is simply that the classical theory usually assumes complete divergence between mainland and island, but that when adaptation is from standing variation instead, this does not hold, since allele frequency differences between mainland and island are only slight. In the latter case, many alleles of small effect can be divergently maintained by selection. Another reason why this does not hold is that selected alleles will be in linkage disequilibrium, so that entire sets of deleterious alleles introduced by migrants are eliminated jointly. The latter phenomenon is explored in detail in Sachdeva (2022) and Zwaenepoel, Sachdeva, and Fraïsse (2024).

I do not think, however, that it is pertinent to discuss these issues in the present paper – they are somewhat technical and require quite some space to explain and do not have much to do with the problem studied, i.e. polyploid establishment, so in the interest of brevity I decided not to revisit this point.

Other comments:

p.1, Abstract: Use the word “found”, but is “establish” more accurate? Overall, the corresponding sentence is difficult to understand/parse. Suggest revision.

This sentence has been revised.

p. 1: “negative frequency dependence” -> Isn’t MCE a positive-frequency dependent process?

Indeed, it is! This has been corrected.

p. 2: Suggest citing Bulmer (1971).

I am unsure whether this is a good idea, as it may lead to some confusion as to which version of the infinitesimal model we are working with. The model described in Bulmer (1971) is not exactly equivalent to the Gaussian descendants approximation that I use (as outlined in Turelli (2017), which I cite). If the reference is suggested by the reviewer in order to cite a more ‘original’ reference, I don’t think there is a reason to single out Bulmer (1971), one could refer to Fisher (1918) or work by Alan Robertson in the 1960’s. I think the reference to Turelli (2017), which reviews the history of the infinitesimal model, is quite adequate, and prevents confusion as to what we mean when we refer to ‘the infinitesimal model’.

p.7: Suggest adding “, across gene copy arrangements and ploidies” at the end of the last sentence before the “Establishment” section.

I do not fully understand this suggestion. I personally find this makes the sentence more confusing, as I never talk of gene copy arrangements.

p. 7, Establishment section: The organization of paragraphs two and three in this section is a bit confusing. N is used before it is formally defined. In my first reading I thought there was migration and then selection, whereas there is selection and then migration. It would also help support a self-contained paper to give the expression for $E[e^{\gamma(\dots)}]$.

As far as I can tell, N is defined when first mentioned. I slightly altered the sentence to make this more clear.

I added the expression for the expected fitness (eq. 11).

The life cycle is indeed as follows: migrants arrive, they mate randomly and offspring survives with a probability that is determined by its fitness. This is indeed ‘selection before migration’. I tried to make this more clear.

p. 11, figure 3 (and later figure 5): The smallest value of m is 0.1 in these figures, but based on figure 4, it looks like $m = 0.01$ is a transition point from processes shaped by the accumulation of inbreeding versus migration becoming more of a factor. Can m be extended to 0.01?

I extended the relevant figures as suggested, as $m = 0.01$ is perhaps more representative of weak migration than $m = 0.1$. Weak migration means essentially that m^2 is negligible, so migrants and early generation descendants thereof do not mate with migrants and the whole process amounts more or less to independent establishment trials (there is no swamping).

I do not think, however, that fig 4 suggests such a transition point though. As noted in the main text the difference between the $m = 0$ and $m = 0.01$ results in fig. 4 are mostly due to the fact that the latter assumes a Gaussian distribution of migrant trait values, whereas the former assumes a fixed initial migrant trait value. (“*The similar trait means for weak migration suggest that establishment depends mostly on the chance pick of a well-adapted migrant (note that the $m = 0$ simulations in fig. 4 assume the initial migrant has $z = 0$, whereas the $m = 0.01$ simulations assume the initial trait value to be Gaussian).*”)

p. 12: Add qualifier “as m increases” at end of last sentence in last full paragraph.

Done.

p. 14, Figure 5 caption: “migration and selfing”

Fixed.

p. 4, suppl., figure S4: “formation”

Fixed.

p. 6, suppl., equation 6: Should the square only be in the denominator?

That’s right, thank you!

p. 10, suppl.: “coancestry coefficients in (10) is “ - missing the “(10)”

Fixed, thank you!

p. 12, suppl.: add “between centromere and locus” after “... recombination happens with probability c ”

Done.

p. 13-15, suppl.: Some notation seemed to not be defined, e.g. $V_{S(2,2)}$, ...

The notation is introduced on p.11 in the beginning of S2.6. I have slightly reorganized the first paragraphs of S2.6 to make the notation and general approach more clear.

References

- Barton, Nicholas H, Alison M Etheridge, and Amandine Véber. 2017. “The Infinitesimal Model: Definition, Derivation, and Implications.” *Theoretical Population Biology* 118: 50–73.
- Barton, Nicholas H, and AM Etheridge. 2018. “Establishment in a New Habitat by Polygenic Adaptation.” *Theoretical Population Biology* 122: 110–27.
- Gallais, André. 2003. “Quantitative Genetics and Breeding Methods in Autopolyploid Plants.”
- Griswold, Cortland K. 2021. “The Effects of Migration Load, Selfing, Inbreeding Depression, and the Genetics of Adaptation on Autotetraploid Versus Diploid Establishment in Peripheral Habitats.” *Evolution* 75 (1): 39–55.
- Oswald, Benjamin P, and Scott L Nuismer. 2011. “A Unified Model of Autopolyploid Establishment and Evolution.” *The American Naturalist* 178 (6): 687–700.
- Porturas, Laura D, Thomas J Anneberg, Anne E Curé, Shengpei Wang, David M Althoff, and Kari A Segraves. 2019. “A Meta-Analysis of Whole Genome Duplication and the Effects on Flowering Traits in Plants.” *American Journal of Botany* 106 (3): 469–76.
- Sachdeva, Himani. 2022. “Reproductive Isolation via Polygenic Local Adaptation in Sub-Divided Populations: Effect of Linkage Disequilibria and Drift.” *PLoS Genetics* 18 (9): e1010297.
- Turelli, Michael. 2017. “Commentary: Fisher’s Infinitesimal Model: A Story for the Ages.” *Theoretical Population Biology* 118: 46–49.

- Walsh, Bruce, and Michael Lynch. 2018. *Evolution and Selection of Quantitative Traits*. Oxford University Press.
- Zwaenepoel, Arthur, Himani Sachdeva, and Christelle Fraïsse. 2024. “The Genetic Architecture of Polygenic Local Adaptation and Its Role in Shaping Barriers to Gene Flow.” *Genetics*, iyae140.