

# Autopolyploid establishment through polygenic adaptation

## Abstract

We define the infinitesimal model of quantitative genetics for the inheritance of an additive quantitative trait in a mixed-ploidy population consisting of diploid, triploid and autotetraploid individuals producing haploid and diploid gametes. We implement efficient simulation methods and use these to study the quantitative genetics of mixed-ploidy populations and the establishment of autotetraploids in a new habitat. We show that, when migration from the source population is rare, autotetraploids are more likely to establish in the new habitat than diploids under a very broad range of conditions, but that this is unlikely to sufficiently counter the scarcity of tetraploid founders when the source is predominantly diploid. We assess in more detail how minority cytotype exclusion interacts with migration load in the establishment process and evaluate the impact of additional sources of prezygotic isolation, specifically selfing and assortative mating, on the relative establishment probabilities of the different cytotypes. In the discussion we consider how inbreeding depression may impact our findings.

**Keywords:** polyploidy, adaptation, establishment, inbreeding, quantitative genetics

## Introduction

Many plant species exhibit ploidy variation (Levin, 2002; Soltis et al., 2007; Rice et al., 2015), and many of these *mixed-ploidy* species have populations in which different cytotypes coexist or form contact zones (Kolář et al., 2017). How such mixed-ploidy populations emerge and are maintained has proven somewhat challenging to understand.

Consider for instance a randomly mating diploid population. Under the commonly accepted view that new polyploid plants are mostly formed through the union of unreduced gametes (Bretagnolle and Thompson, 1995; Herben et al., 2016; Kreiner et al., 2017b), a new tetraploid individual originating by a chance encounter of two unreduced diploid gametes (an event occurring at an appreciable rate; Kreiner et al. (2017a)) is highly unlikely to establish a stable tetraploid subpopulation, as most of its gametes will end up in unfit hybrids of odd ploidy level (a phenomenon referred to as ‘triploid block’, see Ramsey and Schemske (1998); Köhler et al. (2010); Brown et al. (2024)). This positive frequency dependence effect in mixed-ploidy populations is commonly referred to as *minority cytotype exclusion* (MCE), after Levin (1975). It is well-appreciated that, as a consequence of MCE, the rate of unreduced gamete formation needs to be extraordinarily high for tetraploids to establish a stable subpopulation in a large random mating population initially dominated by diploids (Felber and Bever (1997), see also section S2.1).

Hence, to explain the widespread occurrence of mixed-ploidy populations, additional factors besides the continuous formation of polyploids through the union of unreduced

gametes need to be considered. Firstly, chance establishment of tetraploids through drift could occur. Indeed, the problem is somewhat analogous to the spread of underdominant chromosomal rearrangements, where local establishment through genetic drift and subsequent spreading in a subdivided population by means of local extinction and recolonization has been suggested as a plausible model (Lande, 1985). However, MCE is quite strong under random mating, and the population size has to be very small for local tetraploid establishment to occur at an appreciable rate (Rausch and Morgan (2005), see also section S2.2). Secondly, any form of *prezygotic isolation* between cytotypes could promote establishment of polyploid cytotypes by alleviating MCE. Particularly relevant are assortative mating by cytotype (for instance through phenological differences across cytotypes, or differences in pollinators; Husband and Sabara (2004); Kolář et al. (2017)), self-fertilization (Rausch and Morgan, 2005; Novikova et al., 2023), localized dispersal (Baack, 2005; Kolář et al., 2017) and asexual reproduction (Van Drunen and Friedman, 2022). Finally, selection may be invoked to explain the establishment of polyploids. Tetraploids may have higher relative fitness than their diploid counterparts due to reduced inbreeding depression (Husband and Schemske, 1997; Ronfort, 1999; Otto and Whitton, 2000; Husband et al., 2008; Clo and Kolář, 2022), or due to being better adapted to (changing) environmental conditions (Van de Peer et al., 2021). However, none of these factors is likely to explain by itself the establishment of polyploids, and the consensus in the field appears to be that some mix of the above is required to explain the occurrence of mixed-ploidy populations in nature (Kolář et al., 2017; Mortier et al., 2024).

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). This raises the question: which aspects of the process of adaptation to marginal habitats could promote the establishment of polyploid populations? In a peripheral habitat, a new polyploid population may be more likely to reach an appreciable size as it evades to some extent the negative effects of MCE. However, at the same time, such peripheral habitats are likely to present adaptive challenges to establishment (Kawecki, 2008; Sachdeva et al., 2022), and if polyploids are able to colonize such habitats at an appreciable rate, they must somehow be better adapted to local conditions, or more able to adapt to those conditions despite inbreeding and maladaptive migration, compared to diploids. 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.

While there have been substantial modeling efforts aimed at understanding autotetraploid establishment within diploid populations (Levin, 1975; Felber, 1991; Felber and Bever, 1997; Rausch and Morgan, 2005; Oswald and Nuismer, 2011; Clo et al., 2022), the problem of polyploid establishment in peripheral habitats remains largely unaddressed (but see Griswold (2021)), despite its centrality to verbal arguments about the establishment of polyploids in natural populations (Kolář et al., 2017; Van de Peer et al., 2021; Clo, 2022b). Here we develop a model for the establishment of a mixed-ploidy population in a novel, unoccupied habitat based on Barton and Etheridge (2018). In order to estab-

**Table 1:** Glossary of the notation used in the main text.

notation	description
$N$	total population size
$N_k$	population size of the $k$ -ploid cytotype
$\pi_k$	deterministic equilibrium frequency of the $k$ -ploid cytotype
$u$	probability of unreduced gamete formation ( $u = u_{22} = 1 - u_{42}$ )
$v$	probability that a triploid produces a haploid/diploid gamete ( $v = u_{31} = u_{32}$ )
$m$	expected number of migrants per generation arriving in the new habitat
$z_i$	trait value of individual $i$
$c_i$	ploidy level of individual $i$
$g_i$	ploidy level of gamete produced by individual $i$ in a particular cross
$V$	segregation variance in the reference diploid population
$V_{i,k}$	gametic segregation variance associated with the production of a $k$ -ploid gamete by individual $i$
$\mathcal{V}_k$	genetic variance associated with a haploid genome in the $k$ -ploid reference population (i.e. a $k$ -ploid non-inbred population at HWLE)
$\beta_k$	scaling factor for allelic effects in $k$ -ploids
$F_i$	inbreeding coefficient in individual $i$
$\Phi_{ij}$	coancestry coefficient for individuals $i$ and $j$
$\alpha_k$	probability that the two genes at a locus in a diploid gamete formed by a $k$ -ploid individual descend from the same parental gene copy
$\gamma$	strength of directional selection in the new habitat
$\theta$	trait value beyond which the growth rate becomes positive in the new habitat
$w_{ij}$	fitness of parental pair $(i, j)$ (expected fitness of offspring of pair $(i, j)$ )
$w_{ij}^{kl}$	expected fitness of offspring of parental pair $(i, j)$ when $i$ contributes a $k$ -ploid gamete and $j$ contributes a $l$ -ploid gamete
$\sigma_k$	rate of self-fertilization in $k$ -ploids
$\rho_k$	probability of assortative mating in $k$ -ploids

lish in the novel habitat, the population has to adapt to local environmental conditions. We assume fitness is determined by directional selection on a single polygenic trait, which can be interpreted as log fitness at low density in the new habitat. As in Barton and Etheridge (2018), we assume the trait follows the infinitesimal model (*sensu* Barton et al. (2017), i.e. the ‘Gaussian descendants’ infinitesimal model (Turelli, 2017)). We extend the infinitesimal model, and the approach for exact simulation of trait evolution under the infinitesimal model, to mixed-ploidy populations. This is a first contribution of the present paper. We then use simulations to study tetraploid establishment, both from single migrants and under continuous migration from a predominantly diploid source population, examining the effects of autopolyploid genetics, maladaptive migration, selfing and assortative mating on the probability that autotetraploids establish in the novel habitat.

## Model and Methods

### Mixed-ploidy population model

Our notation is summarized in table 1. We consider a mixed-ploidy population of size  $N$  consisting of  $N_2$  diploid,  $N_3$  triploid and  $N_4 = N - N_2 - N_3$  tetraploid individuals. We

105 assume an individual of ploidy level  $k$  forms haploid and diploid gametes with proportions  
 106  $u_{k1}$  and  $u_{k2}$ , as well as a proportion  $1 - u_{k1} - u_{k2}$  inviable (e.g. aneuploid or polyploid)  
 107 gametes. The (relative) fecundity of a  $k$ -ploid individual is hence  $u_{k1} + u_{k2}$ . Unless stated  
 108 otherwise, we will assume

$$\begin{pmatrix} u_{21} & u_{22} \\ u_{31} & u_{32} \\ u_{41} & u_{42} \end{pmatrix} = \begin{pmatrix} 1 - u & u \\ v & v \\ 0 & 1 - u \end{pmatrix} \quad (1)$$

109 where  $u$  is referred to as the proportion of unreduced gametes, and  $2v$  is the proportion  
 110 of euploid gametes produced by a triploid individual. The rate of unreduced gamete  
 111 production is hence fixed across individuals, and is assumed to be the same in diploids  
 112 and tetraploids.

113 When two individuals mate, we assume they produce gametes according to their ploidy  
 114 level (eq. (1)), which randomly combine to produce offspring (which may be inviable if one  
 115 of the contributing gametes is inviable). Intrinsic fitness disadvantages associated with  
 116 particular zygotic ploidy levels or cross types (e.g. modeling phenomena such as ‘triploid  
 117 block’) can be straightforwardly included at this level. An analysis of a deterministic model  
 118 (i.e. where  $N \rightarrow \infty$ ) for the cytotype dynamics and equilibrium cytotype composition  
 119 under random mating is included in section S2.1 (see also Felber and Bever (1997);  
 120 Kauai et al. (2024)). The stochastic version for finite and constant  $N$  is analyzed briefly  
 121 in section S2.2.

## 122 Infinitesimal model

**The basic infinitesimal model.** Consider a population which expresses a quantitative  
 trait determined by a large number of additive loci of small effect. The infinitesimal model  
 approximates the inheritance of such a trait by assuming that the trait value  $Z_{ij}$  of a  
 random offspring from parents with trait values  $z_i$  and  $z_j$  follows a Gaussian distribution  
 with mean equal to the midparent value and variance which is independent of the mean:

$$Z_{ij} \sim \mathcal{N}\left(\frac{z_i + z_j}{2}, V_{ij}\right) \quad (2)$$

123 Here,  $V_{ij}$  is referred to as the *segregation variance* in family  $(i, j)$ . This is the variation  
 124 generated among offspring from the same parental pair due to random Mendelian segre-  
 125 gation in meiosis. This approximation can be justified as arising from the limit where  
 126 the number of loci determining the trait tends to infinity (Barton et al., 2017).

127 An alternative, and for our purposes useful, way to characterize the model is to write  
 128  $Z_{ij} = Y_i + Y_j$ , where  $Y_i$  and  $Y_j$  are independent Gaussian random variables  $Y_i \sim \mathcal{N}\left(\frac{z_i}{2}, V_i\right)$   
 129 (and similarly for  $Y_j$ ). We refer to  $Y_i$  as the (random) *gametic value* of individual  $i$ , and to  
 130  $V_i$  as the *gametic segregation variance* of individual  $i$ . This formulation is helpful in that  
 131 it highlights that Mendelian segregation occurs independently in both parents to produce  
 132 gametes, which then combine additively to determine the offspring trait value. This  
 133 model applies readily to an autoployploid population expressing a trait with infinitesimal  
 134 genetics. However, the segregation variance will be determined by the details of tetraploid  
 135 meiosis, which differ from those of diploid meiosis (see below).

136 In a finite population, the segregation variance will decay over time as the population  
 137 becomes more inbred (Mendelian segregation at homozygous loci does not generate any  
 138 variation). When  $F_i$  is the inbreeding coefficient relative to some ancestral reference

population with gametic segregation variance  $V$  (i.e. the probability that two genes at a locus in individual  $i$  sampled without replacement are identical by descent), the gametic segregation variance of individual  $i$  will be  $V_i = (1 - F_i)V$ . This holds for both diploids and tetraploids (section S2.5.1, also Moody et al. (1993)).

**Scaling of traits across ploidy levels.** If we would naively assume that the allelic effects underlying an additive trait are identical across ploidy levels, a tetraploid offspring from a cross between two diploids would have, on average, a trait value which is the sum of the parental trait values. This is not likely to reflect biological reality: tetraploids do not tend to have, for instance, twice the size of their diploid progenitors on average (e.g. Porturas et al. (2019)). Furthermore, the genetic variance at Hardy-Weinberg and linkage equilibrium (HWLE) in a large non-inbred tetraploid population will be twice that of their diploid counterparts under such assumptions, which is similarly unrealistic (Clo, 2022a).

In order to account for this, we introduce a scaling factor  $\beta_k$ , accounting for the effects of polyploidization *per se* on trait expression in  $k$ -ploids. To introduce and interpret this parameter, we consider an  $L$ -locus additive model, with two alleles (0 and 1) at each locus. For a  $k$ -ploid individual, let  $X_{i,j}$  be the allele at homolog  $j$  of locus  $i$ . We assume the trait value is determined by

$$z = \sum_{i=1}^L \sum_{j=1}^k a_{i,k} X_{i,j} \quad (3)$$

Where  $a_{i,k}$  is the allelic effect of the 1 allele at locus  $i$  in  $k$ -ploids. The genetic variance at HWLE in  $k$ -ploids ( $\tilde{V}_{z,k}$ ) will then be

$$\tilde{V}_{z,k} = k \sum_{i=1}^L a_{i,k}^2 p_i q_i = k \mathcal{V}_k \quad (4)$$

where we refer to  $\mathcal{V}_k$  as the variance associated with a haploid genome in  $k$ -ploids at HWLE. If we now assume  $a_{i,k} = \beta_k a_{i,2}$ , i.e. allelic effects in  $k$ -ploids are as in diploids, but scaled homogeneously by a factor  $\beta_k$ , and assume equal allele frequencies in the different cytotypes, we will have

$$\frac{\tilde{V}_{z,k}}{\tilde{V}_{z,2}} = \frac{k \mathcal{V}_k}{2 \mathcal{V}_2} = \frac{k}{2} \beta_k^2 \quad (5)$$

Note that by definition  $\beta_2 = 1$ . Under the infinitesimal model (where  $a_{i,2} \rightarrow 0$  as  $L \rightarrow \infty$ ), we have  $\tilde{V}_{z,2} = 2 \mathcal{V}_2 = 2V$  (Barton et al., 2017), where  $V$  is the segregation variance in the diploid population. Hence, in the infinitesimal limit we have  $\mathcal{V}_k = \beta_k^2 \mathcal{V}_2 = \beta_k^2 V$ .

**Mixed-ploidy infinitesimal model.** We can extend the infinitesimal model to the mixed-ploidy case, assuming that the gametic value, on the diploid trait scale, associated with a  $k$ -ploid gamete ( $k \in \{1, 2\}$ ) from individual  $i$  of ploidy level  $c_i \in \{2, 3, 4\}$  is a Gaussian random variable  $Y_{i,k}$  with distribution

$$Y_{i,k} \sim \mathcal{N}\left(\frac{k}{c_i} \frac{z_i}{\beta_{c_i}}, V_{i,k}\right) \quad (6)$$

**Table 2:** Gametic segregation variance for haploid and diploid gametes produced by the three cytotypes in the mixed-ploidy model.  $F_i$  is the inbreeding coefficient of individual  $i$  (producing the gamete), whereas  $\alpha_k$  is the probability that a diploid gamete from a  $k$ -ploid individual receives two copies of the same parental gene. We assume  $\alpha_3 \leq 1/4$  and  $\alpha_4 \leq 1/6$  (see section S2.6).

cytotype	haploid gamete variance	diploid gamete variance
diploid	$\frac{1}{2}(1 - F_i)V$	$2\alpha_2(1 - F_i)V$
triploid	$\frac{2}{3}(1 - F_i)V$	$\frac{2}{3}(1 + 3\alpha_3)(1 - F_i)V$
tetraploid	—	$(1 + 2\alpha_4)(1 - F_i)V$

where  $V_{i,k}$  is the gametic segregation variance associated with the production of a  $k$ -ploid gamete by individual  $i$  (see below). The trait value of an individual originating from the union of a  $k$ -ploid gamete of individual  $i$  and an  $l$ -ploid gamete from individual  $j$  is then

$$Z_{ij}^{kl} = \beta_{k+l} (Y_{i,k} + Y_{j,l})$$

i.e.,  $Z_{ij}^{kl}$  is a Gaussian random variate with distribution

$$Z_{ij}^{kl} \sim \mathcal{N} \left( \beta_{k+l} \left( \frac{k}{c_i} \frac{z_i}{\beta_{c_i}} + \frac{l}{c_j} \frac{z_j}{\beta_{c_j}} \right), \beta_{k+l}^2 (V_{i,k} + V_{j,l}) \right) \quad (7)$$

The gametic segregation variance associated with the production of diploid gametes depends not only on the segregation variance in the base population ( $V$ ) and the inbreeding coefficient ( $F$ ), but also on the detailed assumptions of how the meiotic process takes place. Importantly however, the latter only affect the gametic segregation variance through the quantity  $\alpha_k$ , which is the probability that a  $k$ -ploid transmits two copies of the same homolog to a diploid gamete. Note that  $\alpha_4$ , the probability that a diploid gamete of a tetraploid individual carries two copies of the same homolog, is the probability of *double reduction* (e.g. Lynch and Walsh (1998) p.57), and is upper bounded by  $1/6$  (Stift et al., 2008). The value of  $\alpha_2$  depends on the relative frequency of unreduced gamete formation through so-called *first* and *second division restitution* (Bretagnolle and Thompson, 1995; De Storme and Geelen, 2013). We summarize the expressions for the gametic segregation variance in table 2. Detailed derivations can be found in section S2.6.

Writing  $\bar{z}_2$  for the mean trait value in the diploid subpopulation, eq. (7) implies that a tetraploid offspring from a random diploid parental pair has an expected trait value equal to  $\bar{z}_4 = 2\beta_4 \bar{z}_2$ . This hence implies that  $|\bar{z}_4| \geq |\bar{z}_2|$ , with equality only when  $\beta_4 = 1/2$  or  $\bar{z}_2 = \bar{z}_4 = 0$ . In other words, when  $\beta_4 > 1/2$ , we would have for all but  $\bar{z}_2 = 0$  that newly formed tetraploids have more extreme phenotypes on average than their diploid parents. In our establishment model (see below), we shall therefore always consider the case where  $\bar{z}_2 = 0$  in the source population, and think of the trait value modeled as the deviation from the mean phenotype in the source population.

The property that polyploid phenotypes are more extreme on average makes sense if we consider the underlying genetic model: if we assume the source population is at an equilibrium between mutation, drift and stabilizing selection, segregating genetic variants that affect the trait will be equally likely to have positive or negative allelic effects (Hayward and Sella, 2022). At equilibrium, any diploid individual with trait value  $z - \bar{z}_2 > 0$  then carries an excess of variants with positive allelic effects on the trait, and this excess should be exaggerated in its polyploid offspring, which carry *twice* the excess of their parent on average. In other words, the mixed-ploidy model is not ‘coordinate free’ as the basic infinitesimal model is: eq. (7) only makes sense when the trait values that are

200 modeled correspond to deviations from the mean values associated with an underlying  
 201 equilibrium state.

**Recursions for inbreeding coefficients** We can simulate the mixed-ploidy infinitesimal model for a finite population through a straightforward extension of the approach outlined in Barton et al. (2017), provided we can efficiently track inbreeding and coancestry coefficients across the different ploidy levels. Denoting the parents of individual  $i$  by  $k$  and  $l$ , the recursion for the inbreeding coefficients in the mixed-ploidy case becomes

$$\begin{aligned}
 F_i &= \Phi_{kl} & \text{if } c_i &= 2 \\
 F_i &= \frac{1}{3} (F_k^* + 2\Phi_{kl}) & \text{if } c_i &= 3, g_k = 2, g_l = 1 \\
 F_i &= \frac{1}{3} (F_l^* + 2\Phi_{kl}) & \text{if } c_i &= 3, g_k = 1, g_l = 2 \\
 F_i &= \frac{1}{6} (F_k^* + F_l^* + 4\Phi_{kl}) & \text{if } c_i &= 4
 \end{aligned} \tag{8}$$

where  $F_k^* = \alpha_{c_k} + (1 - \alpha_{c_k})F_k$  (section S2.5.1). The recursion for the coancestry coefficients is given by

$$\begin{aligned}
 \Phi_{ii} &= \frac{1}{c_i} (1 + (c_i - 1)F_i) \\
 \Phi_{ij} &= \sum_k \sum_l P_{ik} P_{jl} \Phi_{kl} & i \neq j
 \end{aligned} \tag{9}$$

202 where the sums are over individuals in the parental population, and where  $P_{ik} \in \{0, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, 1\}$   
 203 is the probability that a gene copy in individual  $i$  is derived from parent  $k$ .

## 204 Establishment model

205 Our model for the establishment of a population in an initially unoccupied habitat is  
 206 based on Barton and Etheridge (2018). We assume a large non-inbred ‘mainland’ mixed-  
 207 ploidy population at HWLE and cytotype equilibrium, with  $\mathbb{E}[z] = 0$  irrespective of  
 208 the cytotype. The equilibrium trait value distribution for the different cytotypes on the  
 209 mainland is complicated in general, but a very accurate approximation (which we use  
 210 throughout) is readily obtained for the case where  $u$  is small (see section S2.7).

211 In generation  $t$ ,  $M(t)$  migrant individuals arrive on an island (the new habitat) joining  
 212  $N^*(t)$  resident individuals, where  $M(t)$  is Poisson distributed with mean  $m$ , and  $N^*(0) =$   
 213  $0$  unless stated otherwise. The migrant individuals are assumed to be unrelated to the  
 214 resident individuals. After migration there are  $N(t) = N^*(t) + M(t)$  individuals on the  
 215 island which reproduce sexually, and the offspring thus produced survives until the next  
 216 generation with a probability determined by their trait value. Note that this corresponds  
 217 to a life cycle where selection occurs before migration within a generation. In the basic  
 218 model, random selfing is allowed (but see below for a model with self-incompatibility).  
 219 We assume the trait is under directional selection, with fitness  $w(z) = e^{\gamma(z-\theta)}$ , where  $\gamma$   
 220 is the intensity of directional selection and  $\theta$  is the trait value for which the growth rate  
 221 of the island population becomes positive. As the population becomes better adapted,  
 222 eventually, some form of density regulation must limit its growth. We ignore density  
 223 regulation throughout and focus on the initial establishment phase, defining establishment  
 224 as reaching a population of size 100.

Again following Barton and Etheridge (2018), we simulate the model by first calculating the fitness of each parental pair  $(i, j)$ , which is the expected fitness of offspring of this pair

$$w_{ij} = \sum_{k=1}^2 \sum_{l=1}^2 w_{ij}^{kl} = \sum_{k=1}^2 \sum_{l=1}^2 u_{c_i, k} u_{c_j, l} \mathbb{E} \left[ e^{\gamma(Z_{ij}^{kl} - \theta)} \right] \quad (10)$$

The expectation on the right hand side can be calculated from eq. (7) using the moment-generating function of the Gaussian, i.e.

$$\mathbb{E} \left[ e^{\gamma(Z_{ij}^{kl} - \theta)} \right] = e^{\gamma(\mathbb{E}[Z_{ij}^{kl}] - \theta) + \frac{\gamma^2}{2} \text{Var}[Z_{ij}^{kl}]} \quad (11)$$

Having calculated the  $w_{ij}$ , the number of offspring surviving into the next generation is calculated as  $N^*(t+1) = \sum_{i,j} w_{ij}/N(t)$ . Next,  $N^*(t+1)$  offspring individuals are sampled by sampling parental pairs and gametes proportional to  $w_{ij}^{kl}$ , and sampling a trait value in accordance with eq. (7).

## Self-fertilization and assortative mating

We model partial self-fertilization by assuming that a proportion  $\sigma_{c_i}$  of the ovules of individual  $i$  with ploidy level  $c_i$  are fertilized by self-pollen, while the remaining proportion  $1 - \sigma_{c_i}$  are fertilized by randomly sampled pollen (which may be self-pollen with probability  $1/N$ ). That is, the expected number of offspring from individual  $i$  as mother surviving after selection is

$$\mathbb{E}[w_i] = \sigma_{c_i} w_{ii} + (1 - \sigma_{c_i}) \left[ \frac{1}{N} \sum_{j=1}^N w_{ij} \right] \quad (12)$$

We hence assume no pollen limitation (all outcrossing ovules are fertilized), and no pollen discounting (the probability of being a father is unaffected by an individual's selfing rate). When modeling self-incompatibility, we assume there is no intrinsic disadvantage to self-incompatibility, except when there is only a single individual in the population, i.e.

$$\mathbb{E}[w_i] = \begin{cases} \frac{1}{N-1} \sum_{j \neq i} w_{ij} & \text{if } N > 1 \\ 0 & \text{if } N = 1 \end{cases} \quad (13)$$

We model assortative mating by ploidy level in a similar way, assuming that a fraction  $\rho_{c_i}$  of the ovules of individual  $i$  are fertilized by pollen sampled from the  $c_i$ -ploid portion of the population, while a fraction  $1 - \rho_{c_i}$  is fertilized by pollen randomly sampled from the entire population.

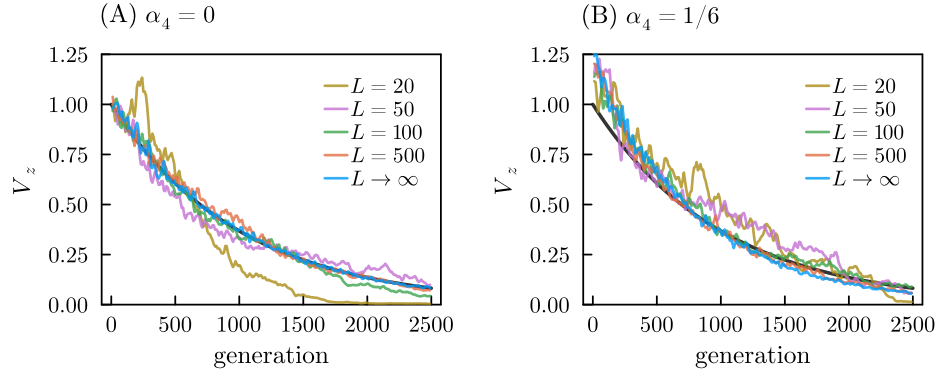
$$\mathbb{E}[w_i] = \rho_{c_i} \frac{1}{N_{c_i}} \sum_{j=1}^N \delta_{c_i, c_j} w_{ij} + (1 - \rho_{c_i}) \left[ \frac{1}{N} \sum_{j=1}^N w_{ij} \right] \quad (14)$$

Where  $\delta_{x,y}$  is the Kronecker delta function.

## Implementation and availability

Individual-based simulations for the mixed-ploidy infinitesimal model were implemented in Julia (Bezanson et al., 2017). Documented code and simulation notebooks are available at [anonymized].





**Figure 1:** The infinitesimal model in autotetraploids. Comparisons are shown for the decay of the genetic variance ( $V_z$ ) due to inbreeding in exact simulations of the infinitesimal model in autotetraploids against individual-based simulations of autotetraploid populations with  $L$  unlinked additive loci determining the quantitative trait. (A) Simulations of a model without double reduction ( $\alpha_4 = 0$ ). (B) Simulations of a model with maximal double reduction ( $\alpha_4 = 1/6$ ) (for all loci in the finite  $L$  simulations). We show window-smoothed values for visual clarity, with observed genetic variances averaged in windows of 20 generations every 10 generations. The black line marks  $e^{-t/4N}$ . We assume  $N = 250$  and  $V_z(0) = 1$ .

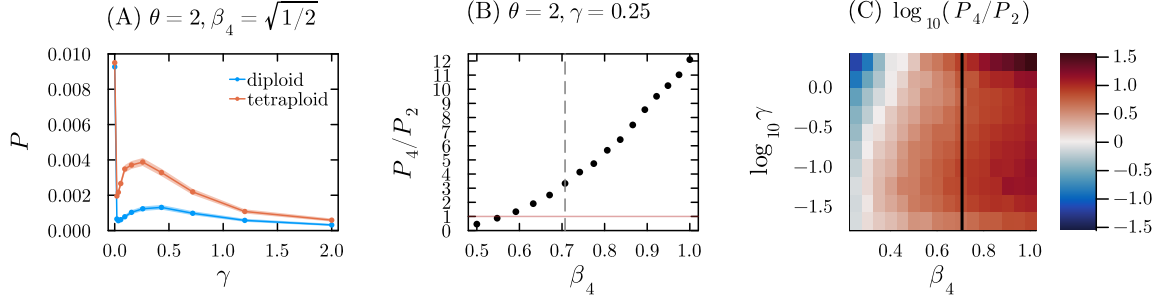
## Results

The results section is organized as follows: first we verify the correctness and accuracy of the infinitesimal model for mixed-ploidy populations by comparison against simulations of a model with finitely many loci. Next we assess how autotetraploid genetics and different assumptions on the genetic variance of neotetraploids affect the probability of tetraploid establishment in a marginal habitat relative to diploid establishment, starting from a single maladapted migrant individual. We then study how continuous migration from a predominantly diploid maladapted source population affects the relative establishment probability, assessing the impact of migration load and MCE. Finally, we consider the impact of prezygotic isolation mechanisms on tetraploid establishment in the marginal habitat. Throughout, we assume  $\beta_3 = \sqrt{2/3}$ ,  $\beta_4 = \sqrt{1/2}$  and  $\alpha_2 = \alpha_3 = \alpha_4 = 0$ , unless stated otherwise. Note that while the former assumption is a natural default, the latter is not easy to motivate, and we do investigate the impact of the  $\alpha$  parameters in some detail below.

### Autotetraploid and mixed-ploidy infinitesimal model

We evaluate the accuracy of the autotetraploid infinitesimal model as an approximation to the evolution of a quantitative trait determined by  $L$  additive loci. We find that the infinitesimal model with inbreeding generally yields accurate predictions for the evolution of the genetic variance when the number of loci is sufficiently large ( $L \geq 100$ , say, figs. 1 and S1). Furthermore, we confirm that, in the absence of double reduction, the decay in genetic variance due to inbreeding after a time  $t$  is well-predicted by  $e^{-t/4N}$  (fig. 1A), as expected from the results of Arnold et al. (2012). As predicted, double reduction (i.e.  $\alpha_4 > 0$ ) leads to an immediate increase in genetic variance (as it increases the segregation variance), but leads to accelerated inbreeding, causing faster loss of variation in the long term (figs. 1 and S1). Simulations for the mixed-ploidy model further confirm the correctness of our infinitesimal approximation (fig. S2).

It is worth noting that, although inbreeding is slower in autotetraploids than in



**Figure 2:** (A) Probability of establishment from a single diploid or tetraploid individual with trait value  $z = 0$  for increasing selection intensity  $\gamma$ . We assume  $m = 0$  and  $u = 0$ , i.e. there is no migration, and no unreduced gametes are produced. The trait is scaled in tetraploids so as to yield the same genetic variance at HWLE ( $\beta_4 = \sqrt{1/2}$ ) (B) Probability of a tetraploid individual with trait value  $z = 0$  successfully founding a population ( $P_4$ ), relative to the probability for a diploid individual with the same trait value ( $P_2$ ). The vertical dashed line marks  $\beta_4 = \sqrt{1/2}$ , for which the variance at HWLE is identical between diploids and autotetraploids. (C) Probability of tetraploid establishment relative to the probability of diploid establishment (on a  $\log_{10}$  scale) across a range of values for  $\gamma$  and  $\beta_4$ . The vertical line again marks  $\beta_4 = \sqrt{1/2}$ . All results are estimated from 500,000 replicate simulations and assume  $\theta = 2$ . Establishment is defined as reaching  $N = 100$ .

diploids for the same population size, the tetraploid fraction of a diploid-dominated mixed-ploidy population will have an equal or higher average inbreeding coefficient (fig. S3). This is because in such a population, triploid and tetraploid individuals mostly arise from gametes formed by diploid individuals, or by polyploid individuals with very recent diploid ancestry (on average  $1 + u + 2v$  generations ago for tetraploids, and  $1 + \frac{2}{3}(u + 2v)$  generations ago for triploids, see section S2.3). A nonzero probability of producing IBD diploid gametes ( $\alpha_k > 0$ ) will then further increase the inbreeding coefficient in the tetraploid and triploid fraction of the population relative to their diploid progenitors (fig. S3). Therefore, as long as diploids dominate, harboring some fraction of the gene pool in polyploid individuals has a negligible effect on the rate of inbreeding in the mixed-ploidy population as a whole, and we find that the evolution of the inbreeding coefficient over time is well predicted by  $1 - e^{-t/2N_e}$ , where the inbreeding-effective population size is, to first order in  $u$ , given by  $(1 - 2u)N$  (section S2.4). This is just the expected number of diploid individuals (to first order in  $u$ ), highlighting that when diploids dominate, polyploids do not contribute to the effective population size.

## Establishment from a single individual

Having established the validity of the mixed-ploidy infinitesimal model, we now use it to study the establishment of polyploids in a marginal habitat to which migrants from a mixed-ploidy source population are maladapted.

We first consider the establishment of a population from a single migrant individual with trait value  $z_0 = 0$ . We assume  $u = 0$  (i.e. there are no unreduced gametes, and hence no newly formed polyploids) and compare the probability of establishment when the migrant is diploid vs. tetraploid (fig. 2). As noted by Barton and Etheridge (2018), the establishment probability (for a given cytotype) depends essentially on two dimensionless parameters,  $\gamma\sqrt{2V}$  and  $\theta/\sqrt{2V}$ , corresponding to the intensity of selection and the degree of maladaptation, respectively. We shall scale our results accordingly, assuming  $2V = 1$  throughout, and shall assume  $\theta = 2$ , so that the average trait value has to increase by two standard deviations to achieve a positive growth rate. The *relative* establishment

probability for different cytotypes will of course also depend on how allelic effects scale across ploidy levels ( $\beta_4$ ).

We find that reduced inbreeding in tetraploids substantially increases the establishment probability of tetraploids relative to diploids across a large part of the parameter space (fig. 2). Indeed, in the case where allelic effects are scaled across ploidy levels to yield the same equilibrium genetic variance ( $\beta_4 = \sqrt{1/2}$ ), the establishment probability for tetraploids can be almost five times as high as for diploids depending on the selection gradient (fig. 2A). As the segregation variance and initial trait value are the same across these simulations, this is a consequence only of the reduced rate of inbreeding, which slows down the exhaustion of the genetic variance carried by the initial migrant individual.

As expected, the probability of establishment decreases when the strength of selection becomes large, as the fitness of the initial migrant individual decreases with increasing  $\gamma$  (fig. 2A). However, when selection becomes very weak, the establishment probability also decreases. This happens because, although the probability of surviving the first couple of generations becomes higher when selection is weaker, adaptation (i.e. increase in the trait mean) will be slower, increasing the risk that genetic variation is exhausted due to inbreeding before the population is able to reach a consistently positive growth rate. Note that when  $\gamma = 0$ , each individual has fitness 1 and we have a critical branching process with a Poisson offspring distribution, so that the probability to reach  $N = 100$  is  $\sim 1/100$  (Barton and Etheridge, 2018).

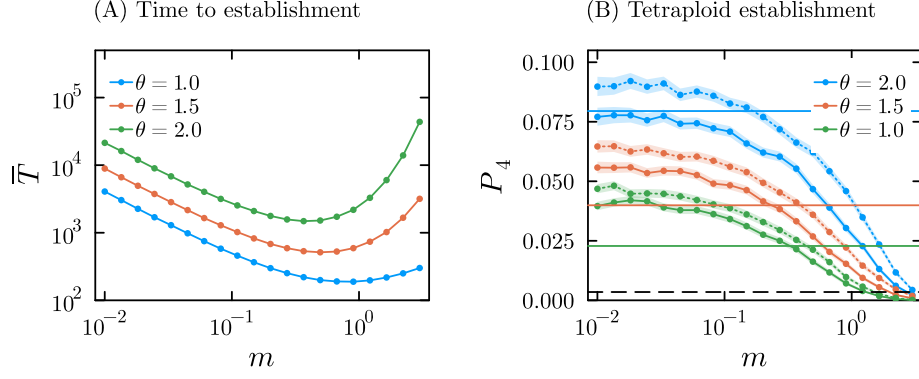
Evidently, the scaling of the genetic variance across ploidy levels has a profound effect on the establishment probability, but only when  $\beta_4$  is close to 0.5 (i.e. individual alleles have almost half the effect size in tetraploids compared to diploids) is the benefit of the slower rate of inbreeding in tetraploids canceled (fig. 2B,C).

## Establishment with recurrent migration

We next consider establishment in the new habitat when there is a continuous influx of migrants ( $m$  migrants per generation on average) coming from a large, noninbred and predominantly diploid source population at cytotype equilibrium. In this setting, establishment is certain to happen eventually, and we are interested in the probability that a tetraploid population establishes before a diploid one does.

We hypothesized that two counteracting processes affect the probability of autotetraploid establishment in this scenario. On the one hand, increased migration will increase the probability that an otherwise likely successful tetraploid migrant suffers from MCE in the early generations while the population size is low, because migrants are likely to be diploid. On the other hand, tetraploids are more strongly reproductively isolated from a typical migrant, so that a tetraploid subpopulation should be less prone to maladaptive gene flow. Hence, conditional on evading MCE, they should be able to adapt to the new habitat at a rate which is not strongly affected by the migration rate. This contrasts with diploids, which interbreed freely with maladapted migrants, resulting in a pulling back of the trait mean towards that of the source population. Lastly, as the mean trait value on the island increases in diploids during adaptation, tetraploid offspring will have more extreme phenotypes on average than diploid offspring when  $\beta_4 > 1/2$ , which may also aid their establishment (irrespective of  $m$ ).

As expected, we find that the time to establishment (of a population of either ploidy level) first decreases with increasing migration as a result of a larger influx of potentially



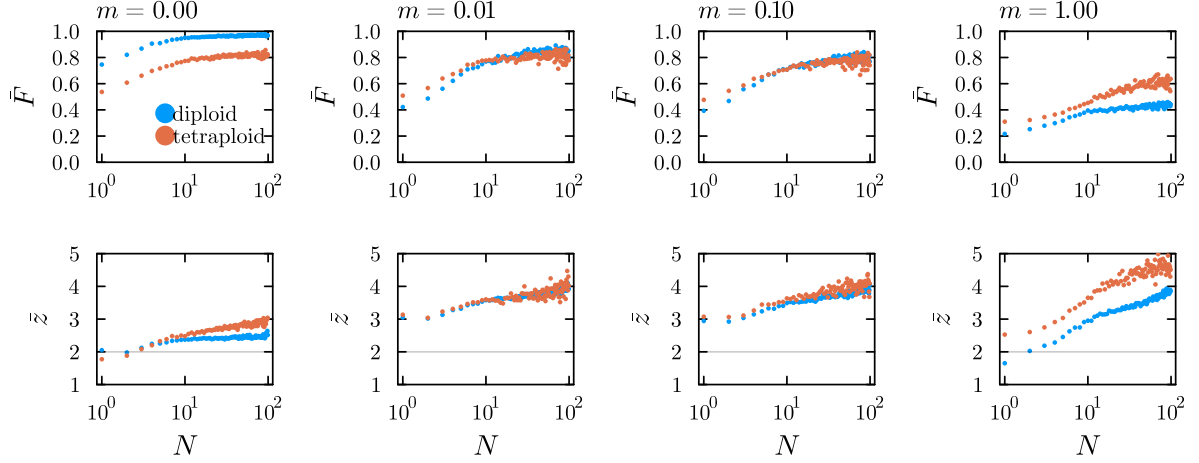
**Figure 3:** Establishment with recurrent migration. (A) Expected time until a population is established in the marginal habitat for increasing rates of migration and different degrees of maladaptation ( $\theta$ ). Results are shown for the case with  $\alpha_k = 0$  for  $k = 2, 3, 4$ . (B) Proportion of simulation replicates in which tetraploids established. The dots connected by solid lines show simulation results with  $\alpha_k = 0$ , whereas the dots connected by dashed lines show simulation results with  $\alpha_2 = 1/2$ ,  $\alpha_3 = 1/4$  and  $\alpha_4 = 1/6$  (i.e. maximum  $\alpha$ ). The solid horizontal lines mark  $\frac{p_4 \pi_4}{(p_4 \pi_4 + p_2 \pi_2)}$ , where  $p_2$  and  $p_4$  are the establishment probabilities for a single diploid and a single tetraploid individual in the absence of migration (as in fig. 2, but with unreduced gamete formation included). The dashed horizontal line marks the proportion of tetraploid migrants (i.e. the proportion of tetraploids at equilibrium in the source population,  $\approx 0.3\%$ ). The baseline predictions (horizontal lines) are based on  $5 \times 10^6$  simulation replicates. All other results are based on 50.000 replicate simulations. We assume  $\gamma = 0.25$  and  $u = v = 0.05$ .

successful migrants, but later increases with increasing migration due to swamping by gene flow (fig. 3A). Importantly, the tetraploid establishment probability is considerably larger than the expected proportion of tetraploid migrants over a large part of the parameter range (fig. 3B, black dashed line). Furthermore, we find that for small migration rates, the probability that tetraploids establish rather than diploids can be larger than expected based on the establishment probabilities in the absence of migration (colored lines in fig. 3), depending on how maladapted migrant individuals are. This suggests that maladaptive gene flow in diploids can increase the probability of tetraploid establishment. However, at the same time we find that the tetraploid establishment probability declines monotonically with  $m$ , which highlights the negative effects of MCE.

Our simulations further show that the mechanism of unreduced gamete formation (as determined by the  $\alpha_2$  parameter) can affect the establishment probability (fig. 3B, dashed lines). This is mainly because the phenotypic variance of a newly formed tetraploid is increased by a factor  $(1 + \alpha_2)$ , thereby increasing the chance that a tetraploid migrant is well-adapted to the marginal habitat. The rate of double reduction ( $\alpha_4$ ) has a more limited effect (fig. S4).

Established diploid populations are more inbred on average than established tetraploids when migration is weak, but the difference is slight except when there is no migration at all (fig. 4, top row). For stronger migration ( $m > 0.1$ ), the opposite holds. This is a result of two interacting processes. On the one hand, inbreeding is slower in tetraploids, so that during adaptation and establishment from a single or limited number of outbred individuals, the inbreeding coefficient is expected to increase less rapidly conditional on the event of tetraploid establishment. On the other hand, migration mostly introduces unrelated diploids, which cross more readily with diploids than tetraploids, reducing the average relatedness more strongly in established diploid than in tetraploid populations.

Conditional on establishment, tetraploids have a higher trait mean both in the absence of migration and for moderate to strong migration, but not when migration is weak

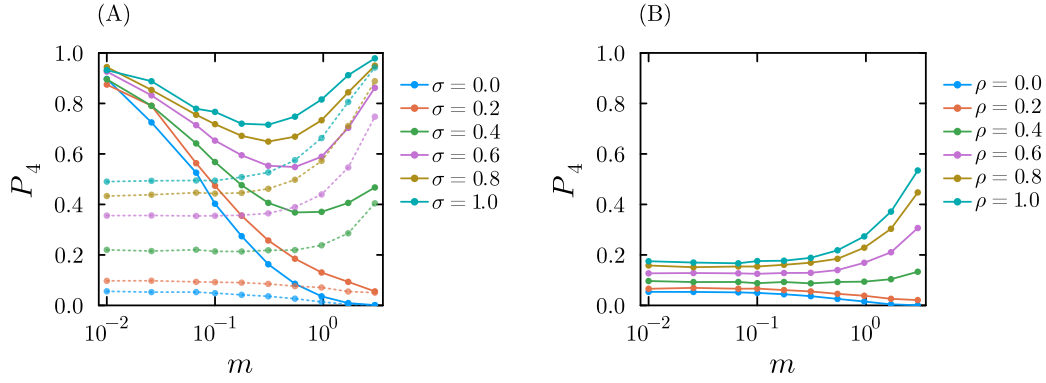


**Figure 4:** Evolution of the mean inbreeding coefficient and trait value across simulation replicates where diploids (blue) or tetraploids (orange) established eventually. Average  $\bar{F}$  and  $\bar{z}$  by population size are shown for increasing rates of migration ( $m$ ) from the predominantly diploid source population. All results are based on 1000 successful establishment replicates. We assume equal equilibrium variance across ploidy levels and  $\gamma = 0.25, \theta = 2, 2V = 1$  and  $u = v = 0.05$ . The  $m = 0$  simulations are seeded with a single diploid or tetraploid individual with trait value 0, and assume  $u = v = 0$ .

(fig. 4, bottom row). In the absence of migration, this is a consequence of the reduced rate of inbreeding and the resulting increased adaptive potential of tetraploids. Another consequence of this is that the average trait value in tetraploids at low population size (conditional on eventual establishment) is lower than that in diploids. This is because, in order to adapt before genetic variation is exhausted, tetraploids need not increase the trait mean as much as diploids do in the first couple of generations. The similar trait means for weak migration suggest that establishment depends mostly on the chance pick of a well-adapted migrant (note that the  $m = 0.01$  simulations assume the initial trait value to be Gaussian). When migration is strong, tetraploids have markedly larger trait values than diploids ( $m = 1$  in fig. 4), showing that diploids suffer strongly from maladaptive gene flow when the population size is low, while tetraploids enjoy some reproductive isolation. Furthermore, in these replicates, tetraploids tend to emerge and rise in frequency at larger population sizes on the island, and hence tend to derive from diploids that already experienced several generations of selection. These neotetraploids, deriving from diploid parents with  $z > 0$ , will have more extreme phenotypes on average (see methods) and hence be better adapted.

### Loss of self-incompatibility, selfing and assortative mating

When polyploidization disrupts an existing SI system (see e.g. Robertson et al. (2011); Zenil-Ferguson et al. (2019); Novikova et al. (2023)), we expect that tetraploids suffer less from MCE, as some portion of their ovules are now assured to be fertilized by diploid gametes, irrespective of the composition of the population. At the same time, we expect that accelerated inbreeding in selfing tetraploids diminishes the adaptive advantage of tetraploids. We find that when polyploidization is associated with the loss of a SI system (i.e. when diploids are self-incompatible, but tetraploids are not), tetraploids have a strongly increased establishment probability (fig. 5). This is the case even when the selfing rate  $\sigma$  in tetraploids is zero (in which case, under our modeling assumptions, there is



**Figure 5:** (A) Establishment with recurrent migration and selfing in polyploids. The solid lines show the case where diploids are self-incompatible. The dashed transparent lines show the case where diploids do random self-fertilization (i.e. self-fertilization occurs with probability  $1/N$ ). Triploids and tetraploids have the same selfing rate.  $\sigma = 0.0$  refers to random self-fertilization. (B) Establishment with recurrent migration and assortative mating by cytotype. The rate of assortative mating is determined by  $\rho_k = \rho$  for  $k = 2, 3, 4$ , where  $\rho_k$  is the probability that an ovule from a  $k$ -ploid mother is pollinated by a  $k$ -ploid father. All results are based on 25,000 replicate simulations. We assume  $\gamma = 0.25, \theta = 1.5$  and  $u = v = 0.05$ .

only random selfing, i.e. the *realized* selfing rate in tetraploids is  $1/N$ ). Furthermore, we find that when the selfing rate is sufficiently high ( $\geq 0.4$  in fig. 5A), the relative establishment probability of tetraploids increases with increasing migration rate. In this regime, the increased strength of MCE due to increased migration and increased reproductive assurance in diploids is compensated by the stronger maladaptive gene flow experienced by diploids.

Self-incompatibility is clearly a strong disadvantage when colonizing a novel habitat, as a self-incompatible population of size one can never reproduce. However, even when diploids are self-compatible, polyploids may still have increased rates of self-fertilization (for instance due to altered flower morphology). For the sake of comparison, fig. 5A also shows results where diploids are assumed to be self-compatible with  $\sigma_2 = 0$  (i.e. random selfing, dashed transparent lines). The tetraploid establishment probability is still markedly increased when  $\sigma_4 \geq 0.4$ , and as for the simulations with self-incompatibility, migration still promotes the probability of tetraploid establishment when the selfing rate in polyploids is sufficiently large compared to the diploid selfing rate.

Another prezygotic isolating mechanism that has often been considered relevant for explaining tetraploid establishment is assortative mating by ploidy level, where ovules from a tetraploid are more likely to be fertilized by pollen coming from a tetraploid – irrespective of the trait values of these individuals. Clearly, assortative mating increases the probability of tetraploid establishment (fig. 5B), although not as strongly as the loss of an SI system does. Again, we find that for some parameter values (roughly  $\rho \geq 0.4$ ), assortative mating may be strong enough so that tetraploid establishment increases with increasing migration rates, suggesting that tetraploids evade maladaptive gene flow sufficiently to overcome MCE. Note that the case  $\rho = 1$  amounts to complete prezygotic isolation.

## Discussion

The observation that polyploid populations tend to inhabit more extreme habitats or occur at the edge of the range of their conspecific diploids has spurred considerable interest among botanists and evolutionary biologists (Kolář et al., 2017; Rice et al., 2019; Van de Peer et al., 2021; Griswold, 2021; Mortier et al., 2024). An important question is whether such patterns emerge because polyploids are somehow more tolerant to extreme environmental conditions (i.e. they somehow are intrinsically more fit than diploids in marginal habitats), or whether other aspects of the population dynamics of mixed-ploidy populations may favor the establishment of polyploid subpopulations.

In this study, we worked out the infinitesimal model for an additive polygenic trait in autotetraploids and mixed-ploidy populations and used it to study the establishment of tetraploids in a marginal habitat by means of individual-based simulations. Assuming the trait to be under directional selection in the marginal habitat, and migration of maladapted individuals from a predominantly diploid source, we sought to determine under which conditions tetraploids are more likely to establish a stable population.

Throughout, 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). We ignore such complications, and hence do not take the actual establishment probabilities very serious, focusing instead on how migration load and prezygotic isolation affect the tetraploid establishment probability.

Similarly, we have ignored mutation, which would reduce the rate at which genetic variation is lost through inbreeding (Barton et al., 2017), and would likely do so differently across cytotypes (i.e.  $\mu V_m$  is expected to differ for different ploidy levels, where  $\mu$  is the mutation rate and  $V_m$  the mutational variance). The contribution of new mutation to the genetic variance on the timescales we consider should however be very limited. Indeed, any individual at the time of establishment derives from a completely outbred migrant individual a relatively short time in the past, so that the opportunity for mutation to contribute to differences in establishment probability between diploids and tetraploids is negligible for realistic  $\mu V_m$ .

Importantly, we assumed no intrinsic advantage or disadvantage of polyploids in the marginal habitat, i.e. the expected fitness of a migrant individual is the same regardless of the ploidy level. Differences in the likelihood of polyploid establishment are hence caused solely by aspects of autopolyploid genetics and the barrier to gene flow between subpopulations of different ploidy levels. This is undoubtedly unrealistic, as both trait values and fitness will often differ systematically across ploidy levels (Porturas et al., 2019). For instance, neopolyploids are likely to suffer intrinsic fertility issues due to meiotic irregularities associated with multivalent formation (Bomblies et al., 2016; Novikova et al., 2023), and triploids may be inviable due to issues with endosperm development (Bretagnolle and Thompson, 1995).

Similarly implausible is the assumption of a constant equilibrium genetic variance across cytotypes ( $\beta_4 = \sqrt{1/2}$  in our model), which we used in most of our results (but see fig. 2). Empirical data on how the genetic variance scales across ploidy levels is scant and suggests that there is no general rule (Gallais, 2003; Porturas et al., 2019). The meta-analysis performed by Porturas et al. (2019) does indicate that trait variance across ploidy levels may often be fairly constant, so our assumption may not yield too bad an approximation. It should be noted that other authors have made different assumptions



on how allelic effects (and hence genetic variance) scale across ploidy levels (in particular Griswold (2021), who scaled allelic effects in a way that is equivalent to  $\beta_4 = 1/2$  in our model). Such assumptions evidently impact the likelihood of polyploid establishment (fig. 2). More empirical data on quantitative traits in experimental or natural mixed-ploidy populations is needed to assess whether the mixed-ploidy infinitesimal model can adequately describe the genetics of quantitative traits across cytotypes, and to suggest plausible values for the relevant parameters ( $\alpha, \beta$ ).

When migration is weak, succesful establishment is not affected by maladaptive gene flow and we can treat establishment in the marginal habitat as independent trials of founding a population from a single individual. In order to avoid extinction, the population has to increase the trait mean by a sufficient amount before the genetic variation carried by the initial migrant individual is exhausted. The probability that the population manages to do so depends on the degree of maladaptation, the intensity of selection and the rate of inbreeding. We find that the decreased rate of inbreeding in autotetraploids gives a rare tetraploid migrant a larger adaptive potential than a diploid migrant, even if the genetic variance carried by the founding individual is the same.

In the presence of maladaptive gene flow, a nascent tetraploid subpopulation suffers from MCE, and although polyploids are more reproductively isolated from a typical migrant (and hence suffer less maladaptive gene flow), MCE will increasingly hamper the establishment of tetraploids as the rate of migration grows. Nevertheless, it is important to remark that despite MCE, the probability of tetraploid establishment in the marginal habitat can be an order of magnitude higher than expected based on the frequency of tetraploid migrants (i.e. is roughly of order  $u$  instead of  $u^2$ ) when migration is sufficiently weak and maladaptation sufficiently high.

Additional sources of prezygotic isolation such as selfing and assortative mating by cytotype may further boost the probability of tetraploid establishment. These processes interact with the rate of migration, so that when selfing/assortative mating occurs above some threshold rate, the tetraploid establishment probability increases with increasing migration rates, whereas below the threshold it decreases with increasing migration pressure. In the latter case, the advantage that tetraploids have when it comes to avoiding maladaptive gene flow is not strong enough to overcome the effects of MCE, whereas in the former case it is.

A major weakness of the present work, and an important caveat, is that we have ignored inbreeding depression and dominance throughout. Including dominance in the infinitesimal framework is already challenging for diploids (requiring the tracking of four-way identity coefficients; Barton et al. (2023)), and appears intractable for higher ploidy levels. However, autopolyploidy has important consequences whenever dominance is relevant, as in the case of inbreeding depression (Ronfort, 1999; Gallais, 2003; Husband et al., 2008; Clo and Kolář, 2022). Indeed, when inbreeding depression is due to recessive deleterious variation, it is expected to be less expressed in neotetraploids because homozygous genotypes should be much rarer than in their diploid parents (the ‘masking’ effect; Husband and Schemske (1997); Otto and Whitton (2000)). Inbreeding during the establishment process should therefore incur a higher fitness cost in diploids relative to neotetraploids, and hence further increase the probability of tetraploid establishment. How this plays out depends however on the *rate* at which populations become inbred, which will differ between cytotypes and will depend strongly on the mating system. In outcrossing populations, inbreeding occurs at a slower rate in tetraploids, further decreasing inbreeding depression and aiding tetraploid establishment. However, when polyploidization is



associated with increased selfing (as when it disrupts an existing SI system), increased inbreeding depression in autotetraploids may prevent their establishment.

Dominance and inbreeding depression may strongly affect the complicated interaction between selfing and migration load in determining tetraploid establishment. Griswold (2021) studied the case where local fitness is determined by a single biallelic locus, and investigated the interaction between inbreeding depression and migration load (where inbreeding depression is modeled as a fixed fitness reduction in offspring produced by selfing). In his model, inbreeding depression is different between cytotypes (assuming stronger inbreeding depression in diploids), so that tetraploids are able to produce more offspring through selfing relative to diploids, who have to rely more on outcrossing. However, outcrossing incurs maladaptive gene flow, and thereby puts the diploids at a disadvantage. He found that autotetraploids can establish when adaptation in the peripheral habitat is conferred by recessive alleles (so that migration load is expressed when migrant alleles are rare) and when inbreeding depression in tetraploids is lower than in diploids. It would be very interesting to combine the infinitesimal framework with some form of inbreeding depression to investigate in a more realistic model whether the combination of maladaptive migration and differential inbreeding depression could explain the prevalence of polyploid subpopulations at range edges.

In the long term, polyploids are expected to accumulate a larger mutation load when deleterious variation is recessive due to less efficient purging, and this may yield *increased* inbreeding depression (Vlček et al., 2025). These effects have been studied in the context of range expansions (Booker and Schrider, 2024). However, this applies only to polyploids that have been established for a long time. In our case, polyploids are always recently descended from diploid ancestors, and they will not have accumulated more deleterious mutations than their diploid counterparts, so that polyploidy should lead to reduced rather than increased inbreeding depression when selfing rates are similar (as discussed above). Interestingly, the interplay between the effects of polyploidy on different timescales could yield an equilibrium situation that may characterize many mixed-ploidy populations in nature: although sometimes polyploids could enjoy enhanced establishment probabilities in peripheral habitats, the accumulation of mutational load may in the long-term limit further range expansion or even lead to competitive exclusion by diploids. Further modeling efforts could provide more insights into the plausibility of such a model.

While in this study we focused on polyploid establishment in a peripheral habitat and how this is affected by migration from a diploid source, the mixed-ploidy infinitesimal framework could be used to address many other eco-evolutionary questions that arise in the study of mixed-ploidy populations. For instance, it could be of interest to develop a complicated individual-based model along the lines of Oswald and Nuismer (2011) to study the effects of selfing, assortative mating and competition on establishment and coexistence of tetraploids within diploid populations, but where the focal trait that determines fitness and assortative mating is not controlled by a few large-effect loci, but many loci of small effect. Similarly, our model could be straightforwardly extended to include population regulation and stabilizing selection, which would allow us to study polyploid establishment along an environmental gradient and the potential of polyploidization to promote range expansions (Polechová and Barton, 2015).

## References

- B. Arnold, K. Bomblies, and J. Wakeley. Extending coalescent theory to autotetraploids. *Genetics*, 192(1):195–204, 2012.
- E. Baack. To succeed globally, disperse locally: effects of local pollen and seed dispersal on tetraploid establishment. *Heredity*, 94(5):538–546, 2005.
- N. Barghi, J. Hermisson, and C. Schlötterer. Polygenic adaptation: a unifying framework to understand positive selection. *Nature Reviews Genetics*, 21(12):769–781, 2020.
- N. H. Barton and A. Etheridge. Establishment in a new habitat by polygenic adaptation. *Theoretical Population Biology*, 122:110–127, 2018.
- N. H. Barton, A. M. Etheridge, and A. Véber. The infinitesimal model: Definition, derivation, and implications. *Theoretical population biology*, 118:50–73, 2017.
- N. H. Barton, A. M. Etheridge, and A. Véber. The infinitesimal model with dominance. *Genetics*, 225(2):iyad133, 2023.
- J. Bezanson, A. Edelman, S. Karpinski, and V. B. Shah. Julia: A fresh approach to numerical computing. *SIAM review*, 59(1):65–98, 2017.
- M. Bohutínská, J. Vlček, S. Yair, B. Laenen, V. Konečná, M. Fracassetti, T. Slotte, and F. Kolář. Genomic basis of parallel adaptation varies with divergence in arabidopsis and its relatives. *Proceedings of the National Academy of Sciences*, 118(21):e2022713118, 2021.
- K. Bomblies and C. L. Peichel. Genetics of adaptation. *Proceedings of the National Academy of Sciences*, 119(30):e2122152119, 2022.
- K. Bomblies, G. Jones, C. Franklin, D. Zickler, and N. Kleckner. The challenge of evolving stable polyploidy: could an increase in “crossover interference distance” play a central role? *Chromosoma*, 125:287–300, 2016.
- W. W. Booker and D. R. Schrider. The genetic consequences of range expansion and its influence on diploidization in polyploids. *The American Naturalist*, 0(ja):null, 2024. doi: 10.1086/733334. URL <https://doi.org/10.1086/733334>.
- F. Bretagnolle and J. D. Thompson. Gametes with the somatic chromosome number: mechanisms of their formation and role in the evolution of autopolyploid plants. *New Phytologist*, 129(1):1–22, 1995.
- M. R. Brown, R. J. Abbott, and A. D. Twyford. The emerging importance of cross-ploidy hybridisation and introgression. *Molecular Ecology*, 33(8):e17315, 2024.
- J. Clo. The evolution of the additive variance of a trait under stabilizing selection after autopolyploidization. *Journal of Evolutionary Biology*, 35(6):891–897, 2022a.
- J. Clo. Polyploidization: Consequences of genome doubling on the evolutionary potential of populations. *American Journal of Botany*, 109(8):1213–1220, 2022b.
- J. Clo and F. Kolář. Inbreeding depression in polyploid species: a meta-analysis. *Biology Letters*, 18(12):20220477, 2022.

- 610 J. Clo, N. Padilla-García, and F. Kolář. Polyploidization as an opportunistic mutation:  
611 The role of unreduced gametes formation and genetic drift in polyploid establishment.  
612 *Journal of Evolutionary Biology*, 35(8):1099–1109, 2022.
- 613 N. De Storme and D. Geelen. Sexual polyploidization in plants—cytological mechanisms  
614 and molecular regulation. *New Phytologist*, 198(3):670–684, 2013.
- 615 F. Felber. Establishment of a tetraploid cytotype in a diploid population: effect of relative  
616 fitness of the cytotypes. *Journal of evolutionary biology*, 4(2):195–207, 1991.
- 617 F. Felber and J. D. Bever. Effect of triploid fitness on the coexistence of diploids and  
618 tetraploids. *Biological Journal of the Linnean Society*, 60(1):95–106, 1997.
- 619 A. Gallais. Quantitative genetics and breeding methods in autopolyploid plants. 2003.
- 620 C. K. Griswold. The effects of migration load, selfing, inbreeding depression, and the  
621 genetics of adaptation on autotetraploid versus diploid establishment in peripheral  
622 habitats. *Evolution*, 75(1):39–55, 2021.
- 623 L. K. Hayward and G. Sella. Polygenic adaptation after a sudden change in environment.  
624 *Elife*, 11:e66697, 2022.
- 625 T. Herben, P. Trávníček, and J. Chrtek. Reduced and unreduced gametes combine  
626 almost freely in a multiploidy system. *Perspectives in Plant Ecology, Evolution and*  
627 *Systematics*, 18:15–22, 2016.
- 628 B. C. Husband and H. A. Sabara. Reproductive isolation between autotetraploids and  
629 their diploid progenitors in fireweed, *chamerion angustifolium* (onagraceae). *New Phy-*  
630 *tologist*, 161(3):703–713, 2004.
- 631 B. C. Husband and D. W. Schemske. The effect of inbreeding in diploid and tetraploid  
632 populations of *epilobium angustifolium* (onagraceae): implications for the genetic basis  
633 of inbreeding depression. *Evolution*, 51(3):737–746, 1997.
- 634 B. C. Husband, B. Ozimec, S. L. Martin, and L. Pollock. Mating consequences of poly-  
635 ploid evolution in flowering plants: current trends and insights from synthetic poly-  
636 ploids. *International journal of plant sciences*, 169(1):195–206, 2008.
- 637 F. Kauai, Q. Bafort, F. Mortier, M. Van Montagu, D. Bonte, and Y. Van de Peer.  
638 Interspecific transfer of genetic information through polyploid bridges. *Proceedings of*  
639 *the National Academy of Sciences*, 121(21):e2400018121, 2024.
- 640 T. J. Kawecki. Adaptation to marginal habitats. *Annual review of ecology, evolution,*  
641 *and systematics*, 39(1):321–342, 2008.
- 642 C. Köhler, O. M. Scheid, and A. Erilova. The impact of the triploid block on the origin  
643 and evolution of polyploid plants. *Trends in Genetics*, 26(3):142–148, 2010.
- 644 F. Kolář, M. Čertner, J. Suda, P. Schönswetter, and B. C. Husband. Mixed-ploidy  
645 species: progress and opportunities in polyploid research. *Trends in plant science*, 22  
646 (12):1041–1055, 2017.

647 V. Konečná, S. Bray, J. Vlček, M. Bohutínská, D. Požárová, R. R. Choudhury,  
648 A. Bollmann-Giolai, P. Flis, D. E. Salt, C. Parisod, et al. Parallel adaptation in au-  
649 topolyploid arabis arenosa is dominated by repeated recruitment of shared alleles.  
650 *Nature Communications*, 12(1):4979, 2021.

651 V. Konečná, M. Šustr, D. Požárová, M. Čertner, A. Krejčová, E. Tylová, and F. Kolář.  
652 Genomic basis and phenotypic manifestation of (non-) parallel serpentine adaptation  
653 in arabis arenosa. *Evolution*, 76(10):2315–2331, 2022.

654 J. M. Kreiner, P. Kron, and B. C. Husband. Frequency and maintenance of unreduced  
655 gametes in natural plant populations: associations with reproductive mode, life history  
656 and genome size. *New Phytologist*, 214(2):879–889, 2017a.

657 J. M. Kreiner, P. Kron, and B. C. Husband. Evolutionary dynamics of unreduced gametes.  
658 *Trends in Genetics*, 33(9):583–593, 2017b.

659 R. Lande. The fixation of chromosomal rearrangements in a subdivided population with  
660 local extinction and colonization. *Heredity*, 54(3):323–332, 1985.

661 D. A. Levin. Minority cytotype exclusion in local plant populations. *Taxon*, 24(1):35–43,  
662 1975.

663 D. A. Levin. *The role of chromosomal change in plant evolution*. Oxford University Press,  
664 USA, 2002.

665 M. Lynch and B. Walsh. *Genetics and analysis of quantitative traits*, volume 1. Sinauer  
666 Sunderland, MA, 1998.

667 M. E. Moody, L. Mueller, and D. Soltis. Genetic variation and random drift in autote-  
668 traploid populations. *Genetics*, 134(2):649–657, 1993.

669 F. Mortier, Q. Bafort, S. Milosavljevic, F. Kauai, L. Prost Boxoen, Y. Van de Peer, and  
670 D. Bonte. Understanding polyploid establishment: temporary persistence or stable  
671 coexistence? *Oikos*, page e09929, 2024.

672 P. Y. Novikova, U. K. Kolesnikova, and A. D. Scott. Ancestral self-compatibility facilitates  
673 the establishment of allopolyploids in brassicaceae. *Plant Reproduction*, 36(1):125–138,  
674 2023.

675 B. P. Oswald and S. L. Nuismer. A unified model of autopolyploid establishment and  
676 evolution. *The American Naturalist*, 178(6):687–700, 2011.

677 S. P. Otto and J. Whitton. Polyploid incidence and evolution. *Annual review of genetics*,  
678 34(1):401–437, 2000.

679 J. Polechová and N. H. Barton. Limits to adaptation along environmental gradients.  
680 *Proceedings of the National Academy of Sciences*, 112(20):6401–6406, 2015.

681 L. D. Porturas, T. J. Anneberg, A. E. Curé, S. Wang, D. M. Althoff, and K. A. Segraves.  
682 A meta-analysis of whole genome duplication and the effects on flowering traits in  
683 plants. *American Journal of Botany*, 106(3):469–476, 2019.

- 684 J. K. Pritchard and A. Di Rienzo. Adaptation—not by sweeps alone. *Nature Reviews*  
685 *Genetics*, 11(10):665–667, 2010.
- 686 J. Ramsey and D. W. Schemske. Pathways, mechanisms, and rates of polyploid formation  
687 in flowering plants. *Annual review of ecology and systematics*, 29(1):467–501, 1998.
- 688 J. H. Rausch and M. T. Morgan. The effect of self-fertilization, inbreeding depression,  
689 and population size on autopolyploid establishmen. *Evolution*, 59(9):1867–1875, 2005.
- 690 A. Rice, L. Glick, S. Abadi, M. Einhorn, N. M. Kopelman, A. Salman-Minkov, J. Mayzel,  
691 O. Chay, and I. Mayrose. The chromosome counts database (ccdb)—a community  
692 resource of plant chromosome numbers. *New Phytologist*, 206(1):19–26, 2015.
- 693 A. Rice, P. Šmarda, M. Novosolov, M. Drori, L. Glick, N. Sabath, S. Meiri, J. Belmaker,  
694 and I. Mayrose. The global biogeography of polyploid plants. *Nature ecology & evolu-*  
695 *tion*, 3(2):265–273, 2019.
- 696 K. Robertson, E. E. Goldberg, and B. Igić. Comparative evidence for the correlated  
697 evolution of polyploidy and self-compatibility in solanaceae. *Evolution*, 65(1):139–155,  
698 2011.
- 699 J. Ronfort. The mutation load under tetrasomic inheritance and its consequences for the  
700 evolution of the selfing rate in autotetraploid species. *Genetics Research*, 74(1):31–42,  
701 1999.
- 702 H. Sachdeva, O. Olusanya, and N. Barton. Genetic load and extinction in peripheral  
703 populations: the roles of migration, drift and demographic stochasticity. *Philosophical*  
704 *Transactions of the Royal Society B*, 377(1846):20210010, 2022.
- 705 D. E. Soltis, P. S. Soltis, D. W. Schemske, J. F. Hancock, J. N. Thompson, B. C. Husband,  
706 and W. S. Judd. Autopolyploidy in angiosperms: have we grossly underestimated the  
707 number of species? *Taxon*, 56(1):13–30, 2007.
- 708 M. Stift, C. Berenos, P. Kuperus, and P. H. van Tienderen. Segregation models for  
709 disomic, tetrasomic and intermediate inheritance in tetraploids: a general procedure  
710 applied to rorippa (yellow cress) microsatellite data. *Genetics*, 179(4):2113–2123, 2008.
- 711 M. Turelli. Commentary: Fisher’s infinitesimal model: A story for the ages. *Theoretical*  
712 *population biology*, 118:46–49, 2017.
- 713 Y. Van de Peer, T.-L. Ashman, P. S. Soltis, and D. E. Soltis. Polyploidy: an evolutionary  
714 and ecological force in stressful times. *The Plant Cell*, 33(1):11–26, 2021.
- 715 W. E. Van Drunen and J. Friedman. Autopolyploid establishment depends on life-history  
716 strategy and the mating outcomes of clonal architecture. *Evolution*, 76(9):1953–1970,  
717 2022.
- 718 J. Vlček, T. Hämälä, C. V. Cobo, E. Curran, G. Šrámková, T. Slotte, R. Schmickl,  
719 L. Yant, and F. Kolář. Whole-genome duplication increases genetic diversity and  
720 load in outcrossing arabidopsis. *bioRxiv*, 2025. doi: 10.1101/2025.01.12.632621. URL  
721 <https://www.biorxiv.org/content/early/2025/01/15/2025.01.12.632621>.

722 R. Zenil-Ferguson, J. G. Burleigh, W. A. Freyman, B. Igić, I. Mayrose, and E. E. Gold-  
723 berg. Interaction among ploidy, breeding system and lineage diversification. *New*  
724 *Phytologist*, 224(3):1252–1265, 2019.