Autopolyploid establishment through polygenic adaptation

Arthur Zwaenepoel[*1]

[1]University of Lille, CNRS, UMR 8198 – Evo-Eco-Paleo, F-59000 Lille, France

Abstract

We define the infinitesimal model of quantitative genetics (*sensu* Barton et al. [2017]) 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.

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 polyploids mostly emerge through the union of unreduced gametes [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 (triploids). This negative frequency dependence effect is commonly referred to as *minority cytotype exclusion* (MCE) after Levin [1975]. It is well-appreciated that, due to MCE, in a large random mating mixed-ploidy population dominated by diploids, the rate of unreduced gamete formation needs to be extraordinarily high for tetraploids to establish (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. On the one hand, chance establishment of tetraploids through drift could occur. However, MCE is quite strong in randomly mating mixed-ploidy populations, and the population size has to be very small for tetraploid establishment to occur at an appreciable rate (Rausch and Morgan [2005], see also section S2.2).

---

[*]arthur.zwaenepoel@univ-lille.fr

On the other hand, 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, [Kolář et al., 2017]), self-fertilization [Rausch and Morgan, 2005], and localized dispersal [Baack, 2005, Kolář et al., 2017]. 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 [Ronfort, 1999], or due to being better adapted to (changing) environmental conditions [Van de Peer et al., 2021].

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]. In particular, polyploids are thought to establish mainly in novel, unoccupied habitats where they evade MCE (for instance at range edges, or after local extinction due to environmental change). If they 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 migration from the source population while the population is still small, compared to diploids. Indeed, when the source population is dominated by diploids, the number of tetraploid migrants arriving in the novel habitat will be almost negligible ($O(u^2)$ if u is the probability that a diploid produces an unreduced gamete in meiosis) compared to the number of diploid migrant individuals, so that tetraploids need a substantial advantage during colonization if they are to establish with reasonable probability.

While there have been substantial modeling efforts aimed at understanding autotetraploid establishment within diploid populations (e.g. Levin [1975], Felber [1991], Felber and Bever [1997], Rausch and Morgan [2005], Oswald and Nuismer [2011], Clo et al. [2022], Griswold [2021]), the problem of polyploid establishment in a novel habitat presenting some adaptive challenge remains largely unaddressed, 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 establish 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). We extend the infinitesimal model, and the approach for exact simulation of trait evolution under the infinitesimal model, to mixed-ploidy populations. 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 polyploid (quantitative) genetics, maladaptive migration, selfing and assortative mating on the probability that autotetraploids establish in the novel habitat before diploids do.

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

notation	description
N, N_k	total population size, population size of the k-ploid cytotype
π_k	equilibrium frequency of the k-ploid cytotype
u	probability of unreduced gamete formation ($u = u_{22} = u_{32} = 1 - u_{42}$)
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
β_k	scaling factor for allelic effects in k-ploids
F_i	inbreeding coefficient in individual i
Φ_{ij}	coancestry coefficient for individuals i and j
α_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
γ	strength of directional selection in the new habitat
θ	trait value beyond which the growth rate becomes positive in the new habitat
w_{ij}	fitness of parental pair (i, j)
w_{ij}^{kl}	expected fitness of offspring from parental pair (i, j) when i contributes a k-ploid gamete and j contributes a l-ploid gamete
σ_k	selfing rate in k-ploids
ρ_k	probability of assortative mating in k-ploids

Model and Methods

Mixed-ploidy population model

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

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

where u is referred to as the proportion of unreduced gametes, and $2v$ is the proportion of euploid gametes produced by a triploid individual.

When two individuals mate, we assume they produce gametes according to their ploidy level (eq. (1)), which randomly combine to produce offspring (which may be inviable if one of the contributing gametes is inviable). Intrinsic fitness disadvantages associated with particular zygotic ploidy levels or cross types (e.g. modeling phenomena such as 'triploid block', [Ramsey and Schemske, 1998, Brown et al., 2024]) can be

straighforwardly included at this level. An analysis of a deterministic model (i.e. where $N \to \infty$) for the cytotype dynamics and equilibrium cytotype composition under random mating is included in section S2.1 (see also Felber and Bever [1997], Kauai et al. [2024]). The stochastic version for finite and constant N is analyzed briefly in section S2.2.

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:

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

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

An alternative, and for our purposes useful, way to characterize the model is to write $Z_{ij} = Y_i + Y_j$, where Y_i and Y_j are independent Gaussian random variables $Y_i \sim \mathcal{N}\left(\frac{z_i}{2}, V_i\right)$ (and similarly for Y_j). We refer to Y_i as the (random) *gametic value* of individual i, and to V_i as the *gametic segregation variance* of individual i. This formulation is helpful in that it highlights that Mendelian segregation occurs independently in both parents when gametes are produced, which then combine additively to determine the offspring trait value. This model applies readily to an autopolyploid population expressing a trait with infinitesimal genetics. For instance, when an autotetraploid produces a diploid gamete, it will pass on half its trait value to the gamete on average, with a variance determined by the details of tetraploid meiosis (which are not, in general, identical to those of diploid meiosis, see below).

Note that in a finite population, the segregation variance will decay over time as the population becomes more inbred. Indeed, Mendelian segregation generates less variation among the gametes produced by individual i when that individual is more inbred, as segregation at homozygous loci does not generate any 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]).

Mixed-ploidy infinitesimal model. We can extend this model to the mixed-ploidy case, where we assume that the gametic value associated with a k-ploid gamete ($k \in$

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 α_k is the probability that a diploid gamete from a k-ploid individual receives two copies of the same parental gene. Note that we have $\alpha_3 \leq 1/4$ and $\alpha_4 \leq 1/6$.

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	$-$	$(1 + 2\alpha_4)(1 - F_i)V$

$\{1, 2\}$) from individual i of ploidy level $c_i \in \{2, 3, 4\}$ is a Gaussian random variable with distribution

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

where $V_{i,k}$ is the gametic segregation variance associated with the production of a k-ploid gamete by individual i. For diploids producing diploid unreduced gametes and triploid producing euploid gametes, the latter depends not only on the segregation variance in the base population (V) and the inbreeding coefficient (F), but also on the detailed assumptions on how these aberrant meiotic processes occur. Importantly however, when only haploid and diploid gametes are assumed to occur, these details only affect the gametic segregation variance through the quantity α_k, which is the probability that a k-ploid transmits two copies of the same homolog to a diploid gamete.

Note that α_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 α_2 depends on the relative frequency of unreduced gamete formation through so-called *first* and *second division restitution* [Bretagnolle et al., 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 S2.6.2.

If we would somewhat naively assume $Z_{ij} = Y_{i,k} + Y_{j,l}$ for a $(k+l)$-ploid offspring from parental pair (i, j), as we did in the standard infinitesimal model (eq. (2)), a tetraploid offspring from a cross between two diploids will yield, on average, a trait value which is the sum of the parental trait values (and hence twice the midparent value). 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. Similarly, the genetic variance at Hardy-Weinberg and linkage equilibrium (HWLE) in tetraploids 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 β_k, accounting for the effects of polyploidization *per se* on trait expression. Specifically, we assume that β_k determines how the variance $\tilde{V}_{z,k}$ at Hardy-Weinberg and linkage equilibrium (HWLE)

in k-ploids is related to that in their diploid counterparts:

$$\frac{\tilde{V}_{z,k}}{\tilde{V}_{z,2}} = \frac{k}{2}\beta_k^2 \tag{3}$$

In the discrete locus model, β_k is interpreted as a scaler of additive effects in the cytotype of ploidy level k (see section S2.6.1 for details). Under these assumptions, and in the infinitesimal limit, the trait value of a k-ploid descendant from parental pair i and j, where parent i contributes a gamete of ploidy level g_i and likewise parent j a gamete of ploidy level g_j, is distributed according to a Gaussian:

$$Z_{ij} \sim \mathcal{N}\left(\beta_k\left(\frac{g_i}{c_i}\frac{z_i}{\beta_{c_i}} + \frac{g_j}{c_j}\frac{z_j}{\beta_{c_j}}\right), \beta_k^2(V_{i,g_i} + V_{j,g_j})\right) \tag{4}$$

where the values of V_{i,g_i} are as in table 2 (see section S2.6.1).

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

$$
\begin{aligned}
F_i &= \Phi_{kl} && \text{if } c_i = 2 \\
F_i &= \frac{1}{3}\left(F_k^* + 2\Phi_{kl}\right) && \text{if } c_i = 3, g_k = 2, g_l = 1 \\
F_i &= \frac{1}{3}\left(F_l^* + 2\Phi_{kl}\right) && \text{if } c_i = 3, g_k = 1, g_l = 2 \\
F_i &= \frac{1}{6}(F_k^* + F_l^* + 4\Phi_{kl}) && \text{if } c_i = 4
\end{aligned}
\tag{5}
$$

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

$$
\begin{aligned}
\Phi_{ii} &= \frac{1}{c_i}\left(1 + (c_i - 1)F_i\right) \\
\Phi_{ij} &= \sum_k \sum_l P_{ik}P_{jl}\Phi_{kl} && i \neq j
\end{aligned}
\tag{6}
$$

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

Establishment model

Our model for the establishment of a population in an initially unoccupied habitat is based on Barton and Etheridge [2018]. We assume a large non-inbred 'mainland' population at HWLE and cytotype equilibrium, with $\mathbb{E}[z] = 0$ irrespective of the cytotype.

In generation t, $M(t)$ migrant individuals arrive on an island (the new habitat) joining $N^*(t)$ resident individuals, where $M(t)$ is Poisson distributed with mean m. The migrant individuals are assumed to be unrelated to the resident individuals. After migration in generation t, the $N(t) = N^*(t) + M(t)$ individuals reproduce sexually, and the offspring thus produced survives until the next generation with a probability determined by their trait value. In the basic model, random selfing is allowed (but see below for a model with self-incompatibility). We assume the trait is under directional selection, with fitness is $w(z) = e^{\gamma(z-\theta)}$, where γ is the intensity of directional selection and θ is the trait value for which the growth rate of the population becomes positive.

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

$$w_{ij} = \sum_{k=1}^{2} \sum_{l=1}^{2} w_{ij}^{kl} = \sum_{k=1}^{2} \sum_{l=1}^{2} u_{c_i,k} u_{c_j,l} \mathbb{E}\left[e^{\gamma(Z_{ij}-\theta)} | g_i = k, g_j = l\right] \tag{7}$$

The expectation on the right hand side can be calculated from eq. (4) using the moment-generating function of the Gaussian. 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. (4).

Self-fertilization and assortative mating

We model partial self-fertilization by assuming that a proportion σ_{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

$$\mathbb{E}[w_i] = \sigma_{c_i} \mathbb{E}[w_{ii}] + (1 - \sigma_{c_i}) \left[\frac{1}{N} \sum_{j=1}^{N} \mathbb{E}[w_{ij}]\right] \tag{8}$$

Where $\mathbb{E}[w_{ij}] = e^{\gamma \bar{z}_{ij} + \gamma^2 V_{ij}/2}$ is the expected fitness of offspring from the parental pair with i as mother and j as father. 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.

$$\mathbb{E}[w_i] = \begin{cases} \frac{1}{N-1} \sum_{j \neq i} \mathbb{E}[w_{ij}] & \text{if } N > 1 \\ 0 & \text{if } N = 1 \end{cases} \tag{9}$$

We model assortative mating by ploidy level in a similar way, assuming that a fraction ρ_{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.

$$\mathbb{E}[w_i] = \sum_{j=1}^{N} \left(\frac{\delta_{c_i,c_j}}{N_{c_i}} \rho_{c_i} + \frac{1}{N}(1 - \rho_{c_i}) \right) \mathbb{E}[w_{ij}] \qquad (10)$$

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 `https://github.com/arzwa/InfGenetics`.

Results

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 leads to an immediate increase in genetic variance, but leads to accelerated inbreeding, causing faster loss of genetic variance in the long-term (figs. 1 and S1). As noted by Arnold et al. [2012], the loss of genetic variation in the presence of double reduction cannot, however, be described by a single long-term effective population size.

Simulations for the mixed-ploidy model further confirm the correctness of our infinitesimal approximation and highlight the importance of assumptions regarding the scaling of allelic effects across ploidy levels (β parameters, fig. S2). Although inbreeding is slower in autotetraploids than in 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 polyploid individuals with very recent diploid ancestry (about $1 + u + 2v$ generations ago for tetraploids, and $1 + \frac{2}{3}(u + 2v)$ generations ago for triploids, see section S2.3), so that the polyploid subpopulations will show an average relatedness similar to the diploid population and not evolve as an isolated higher-ploidy population would. 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, the effect of harboring some of the gene pool in triploid and tetraploid individuals on the rate of inbreeding is negligible, and we find that the evolution of the inbreeding coefficient over
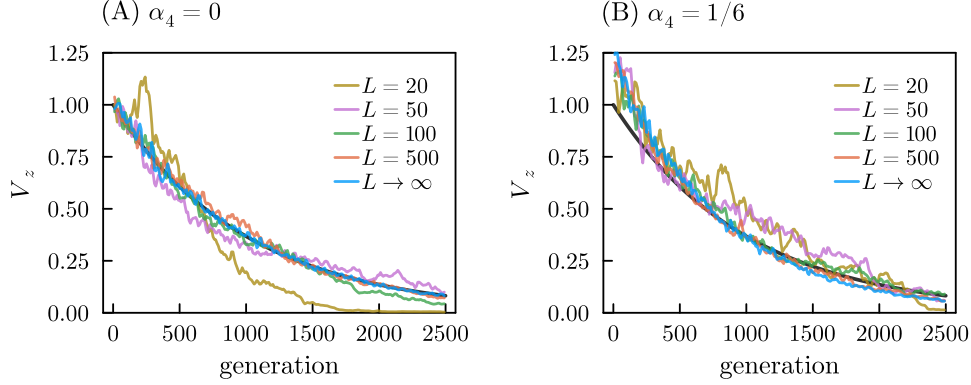
8

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 $V_z(0) = 1$.

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).

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 $u = 0$ 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 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.

We find that reduced inbreeding in tetraploids substantially increases the establishment probability. Indeed, in the case where allelic effects are scaled so as to yield the same equilibrium genetic variance ($\beta = \sqrt{1/2}$), the establishment probability for tetraploids can be almost five times as high as in diploids depending on the selection gradient (fig. 2A). While the probability of establishment becomes lower as the strength of selection becomes large, the establishment probability does not decrease monotonically with γ, i.e. as selection becomes very weak, the establishment probability also lowers. This happens because, although the probability of surviving the first couple of generations becomes higher for weaker selection, adaptation will be slower, and the risk
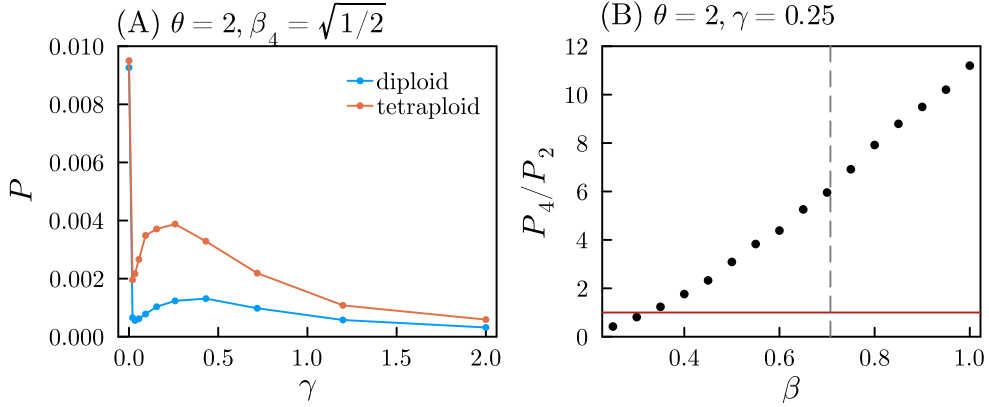
Figure 2: (A) Probability of establishment from a single diploid or tetraploid individual with trait value $z = 0$ for increasing selection intensity γ. We assume $m = 0$ and $u = 0$, i.e. there is no migration, and no unreduced gametes are produced. The trait is scaled in tetraploids so as to yield 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. All results are estimated from 500.000 replicate simulations.

that genetic variation is exhausted due to inbreeding before the population is able to reach a consistently positive growth rate is increased.

As expected, the scaling of the equilibrium genetic variance has a profound effect on the establishment probability, but only when β 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).

Establishment under recurrent migration

We next consider establishment in the new habitat when there is a continuous influx of migrants coming from a large diploid-dominated mixed-ploidy source population at cytotype equilibrium. In this case, establishment is certain to happen eventually, and we are interested in the probability that a tetraploid population established before a diploid one does.

We hypothesized that two counteracting processes affect the probability of autotetraploid establishment in this case. 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 strongly reproductively isolated from the typical migrant, so that they are less prone to maladaptive swamping. 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.
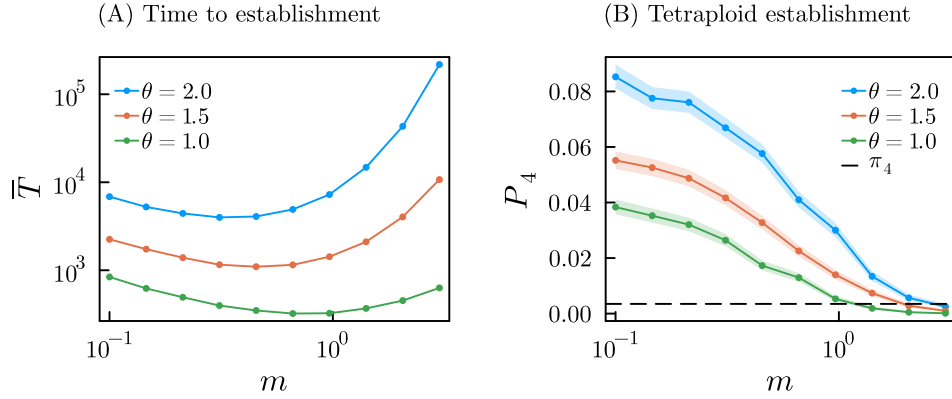
(A) Time to establishment   (B) Tetraploid establishment

Figure 3: Establishment with recurrent migration. (A) Expected time until a population is established in the novel habitat for increasing rates of migration and different degrees of maladaptation ($\bar{z}_s$, the mean trait value in the source population). (B) Proportion of simulation replicates in which tetraploids established. The shaded area corresponds to Jeffreys' 95% interval. The horizontal dashed line marks the proportion of tetraploid migrants (i.e. the proportion of tetraploids at equilibrium in the source population). All results are based on 10.000 replicate simulations. We assume $\gamma = 0.25$ and $u = v = 0.05$.

Loss of self-incompatibility

Assortative mating

Discussion

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. 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.

J. Bezanson, A. Edelman, S. Karpinski, and V. B. Shah. Julia: A fresh approach to numerical computing. *SIAM review*, 59(1):65–98, 2017.

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.
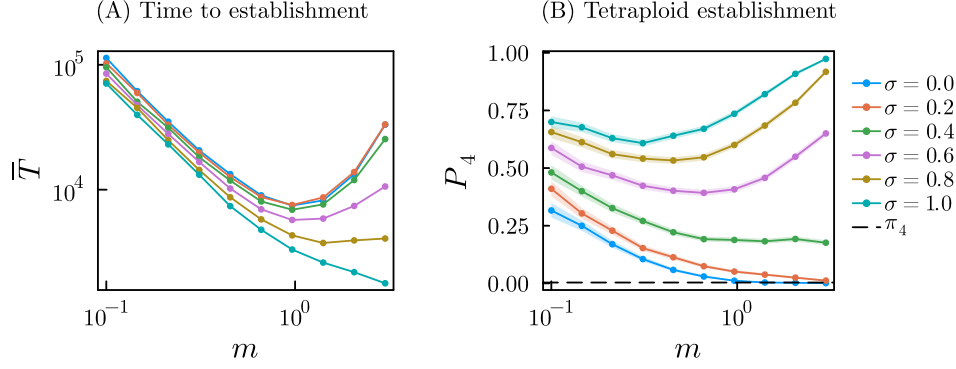
(A) Time to establishment (B) Tetraploid establishment

Figure 4: Establishment with recurrent migration and loss of self-incompatibility in polyploids. (A) Expected time until a population is established in the novel habitat for increasing rates of migration and different self-fertilization rates in polyploids. Diploids are assumed to be self-incompatible. Note that $\sigma = 0.0$ refers to random self-fertilization (i.e. self-fertilization with probability $1/N$). (B) Proportion of simulation replicates in which tetraploids established. The shaded area corresponds to Jeffreys' 95% interval. All results are based on 10.000 replicate simulations. We assume $\gamma = 0.25$ and $u = v = 0.05$.

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, N. Padilla-García, and F. Kolář. Polyploidization as an opportunistic mutation:
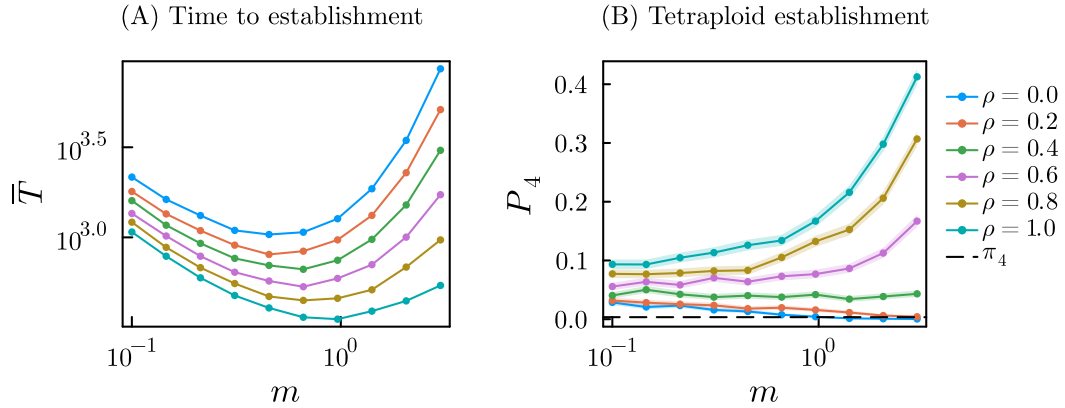
(A) Time to establishment (B) Tetraploid establishment

Figure 5

The role of unreduced gametes formation and genetic drift in polyploid establishment. *Journal of Evolutionary Biology*, 35(8):1099–1109, 2022.

N. De Storme and D. Geelen. Sexual polyploidization in plants–cytological mechanisms and molecular regulation. *New Phytologist*, 198(3):670–684, 2013.

F. Felber. Establishment of a tetraploid cytotype in a diploid population: effect of relative fitness of the cytotypes. *Journal of evolutionary biology*, 4(2):195–207, 1991.

F. Felber and J. D. Bever. Effect of triploid fitness on the coexistence of diploids and tetraploids. *Biological Journal of the Linnean Society*, 60(1):95–106, 1997.

C. K. Griswold. 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, 2021.

T. Herben, P. Trávníček, and J. Chrtek. Reduced and unreduced gametes combine almost freely in a multiploidy system. *Perspectives in Plant Ecology, Evolution and Systematics*, 18:15–22, 2016.

F. Kauai, Q. Bafort, F. Mortier, M. Van Montagu, D. Bonte, and Y. Van de Peer. Interspecific transfer of genetic information through polyploid bridges. *Proceedings of the National Academy of Sciences*, 121(21):e2400018121, 2024.

F. Kolář, M. Čertner, J. Suda, P. Schönswetter, and B. C. Husband. Mixed-ploidy species: progress and opportunities in polyploid research. *Trends in plant science*, 22 (12):1041–1055, 2017.

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

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

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

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

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

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

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

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

J. Ramsey and D. W. Schemske. Pathways, mechanisms, and rates of polyploid formation in flowering plants. *Annual review of ecology and systematics*, 29(1):467–501, 1998.

J. H. Rausch and M. T. Morgan. The effect of self-fertilization, inbreeding depression, and population size on autopolyploid establishmen. *Evolution*, 59(9):1867–1875, 2005.

A. Rice, L. Glick, S. Abadi, M. Einhorn, N. M. Kopelman, A. Salman-Minkov, J. Mayzel, O. Chay, and I. Mayrose. The chromosome counts database (ccdb)–a community resource of plant chromosome numbers. *New Phytologist*, 206(1):19–26, 2015.

J. Ronfort. The mutation load under tetrasomic inheritance and its consequences for the evolution of the selfing rate in autotetraploid species. *Genetics Research*, 74(1): 31–42, 1999.

D. E. Soltis, P. S. Soltis, D. W. Schemske, J. F. Hancock, J. N. Thompson, B. C. Husband, and W. S. Judd. Autopolyploidy in angiosperms: have we grossly underestimated the number of species? *Taxon*, 56(1):13–30, 2007.

M. Stift, C. Berenos, P. Kuperus, and P. H. van Tienderen. Segregation models for disomic, tetrasomic and intermediate inheritance in tetraploids: a general procedure applied to rorippa (yellow cress) microsatellite data. *Genetics*, 179(4):2113–2123, 2008.

Y. Van de Peer, T.-L. Ashman, P. S. Soltis, and D. E. Soltis. Polyploidy: an evolutionary and ecological force in stressful times. *The Plant Cell*, 33(1):11–26, 2021.