Polyploidization as an opportunistic mutation: The role of unreduced gametes formation and genetic drift in polyploid establishment

Josselin Clo, Nélida Padilla-García, Filip Kolář

Polyploidization as an opportunistic mutation: the role of unreduced gametes formation and genetic drift in polyploid establishment

Josselin Clo [1], Nélida Padilla-García [2], Filip Kolář [3]

[1]Department of Botany, Faculty of Science, Charles University in Prague, Benátská 2, CZ-128 01 Prague, Czech Republic, ORCID **0000-0002-3295-9481**, josselin.clo@gmail.com

[2]Department of Botany, Faculty of Science, Charles University in Prague, Benátská 2, CZ-128 01 Prague, Czech Republic, ORCID **0000-0002-4877-6743**, nelidapg@gmail.com

[3]Institute of Botany of the Czech Academy of Sciences, Zámek 1, CZ-252 43 Průhonice, Czech Republic and Department of Botany, Faculty of Science, Charles University in Prague, Benátská 2, CZ-128 01 Prague, Czech Republic, ORCID **0000-0002-8793-7992**, filip.kolar@gmail.com

Contribution: J.C. initiated the project, developed the mathematical and simulation models and wrote the initial draft of the manuscript. J.C. N. P-G. and F.K. edited the manuscript.

Acknowledgment: We thank P. Meirmans for interesting discussions during the project, and C. Fraïsse for valuable comments on a previous version of the manuscript. We thank L. Holman and two anonymous reviewers for valuable comments that improved the quality of the manuscript. This work was supported by the European Research Council (project 850852 DOUBLEADAPT). Additional support was provided by the Czech Science Foundation (project 20-22783S) and the Czech Academy of Sciences (long-term research development project no. RVO 67985939).

Conflicts of interest: We have nothing to declare.

Data availability: The simulation code has been deposited on Github (https://github.com/JosselinCLO/Drift_polyploidy).

Abstract: It is broadly assumed that polyploidy success reflects an increase in fitness associated with whole-genome duplication (WGD), due to higher tolerance to stressful conditions. Nevertheless, WGD also arises with several costs in neo-polyploid lineages, like genomic instability, or cellular mis-management. In addition to these costs, neo-polyploid individuals also face frequency dependent selection, because of frequent low-fitness triploids formed by cross-ploidy pollinations when tetraploids are primarily rare in the population. Interestingly, the idea that polyploidy can be fixed by genetic drift, as a neutral or deleterious mutation is currently underexplored in the literature. To test how and when polyploidy can fix in a population by chance, we built a theoretical model in which autopolyploidization occurs through the production of unreduced gametes, a trait modeled as a quantitative trait that is allowed to vary through time. We found that when tetraploid individuals are less or as fit as their diploid progenitors, fixation of polyploidy is only possible when genetic drift is stronger than natural selection. The necessity of drift for tetraploid fixation holds even when polyploidy confers a selective advantage, except for scenarios where tetraploids are much fitter than diploids. Finally, we found that self-fertilization is less beneficial for tetraploid establishment than previously thought, notably when polyploids harbour an initial decrease in fitness. Our results bring a novel, non-exclusive explanation for the unequal temporal and spatial distribution of polyploid species.

Keywords: Polyploid establishment, self-fertilization, unreduced gametes, genetic drift

INTRODUCTION

Whole-genome duplication is present throughout the evolution of eukaryotes (Mable, 2004; Gregory & Mable, 2005), with a strong representation in angiosperms for which 15% of speciation events are associated with a ploidy increase (Ramsey & Schemske, 2002; Soltis *et al.*, 2007; Wood *et al.*, 2009; Parisod *et al.*, 2010; Barker *et al.*, 2016). Polyploidy has been demonstrated to have a broad range of effects on plant phenotypes and genomes. Polyploidization is hypothesized to be both an evolutionary dead-end (Stebbins, 1971; Mayrose *et al.*, 2011), and an important driver of plant adaptation and speciation (Van de Peer *et al.*, 2017). Notably, the apparent non-random establishment probability of genome doubling events has suggested a link with environmental conditions (Van de Peer *et al.*, 2021). Polyploidization is known to confer a selective advantage under several stressful conditions (Bomblies, 2020; Van de Peer *et al.*, 2021). For example, polyploidy has been associated with higher tolerance to salt (Chao *et al.*, 2013) and drought stress (Ruiz *et al.*, 2016), and higher resistance to pathogens and herbivory (Hannweg *et al.*, 2016; Hias *et al.*, 2018; Wang *et al.*, 2018). All these advantages tend to explain why successful polyploidization events are associated with periods of dramatic global change (Vanneste *et al.*, 2014a), or harsher environments (Rice *et al.*, 2019).

As shown before, polyploidization can be favored during unstable and global change periods. Nevertheless, it has been suggested that genome doubling should be counter-selected in stable periods and environments (Comai, 2005; Oberlander *et al.*, 2016; Kreiner *et al.*, 2017a; Van de Peer *et al.*, 2017). First, because polyploidization arises with several costs in neo-polyploid lineages, like genomic instability, mitotic and meiotic abnormalities, and a reduction in fitness (Comai, 2005; Otto, 2007; Doyle & Coate, 2019; Porturas *et al.*, 2019;

Clo & Kolář, 2021). In addition to these intrinsic costs of polyploidization, neo-polyploid individuals are also facing frequency dependent selection. Newly formed polyploid individuals are primarily in minority within populations, and most of their reproductive events with diploid individuals lead to low-fitness triploids, enhancing the extinction probability of polyploids (a phenomenon called Minority Cytotype Exclusion, Levin, 1975). Second, the production of polyploid individuals is often a costly mechanism for the diploid ancestral populations (Kreiner *et al.*, 2017a). It is recognized that the production of unreduced gametes is the primary mechanism of polyploid formation (Harlan & deWet, 1975; Ramsey & Schemske, 1998; Tayalé & Parisod, 2013; Mason & Pires, 2015). The production of unreduced gametes is expected to be counter-selected in stable environments, because for each unreduced gamete formed, two normal gametes are lost (Kreiner *et al.*, 2017a), which can potentially reduce the contribution of an individual to fertilization events. In addition, if the production of unreduced gametes is low in a population, most fertilization events involving unreduced gametes will lead to the production of triploid individuals, which generally harbor low fitness (Ramsey & Schemske, 1998; Burton & Husband, 2000; Husband & Sabara, 2003). The expectation of low unreduced gamete production in natural populations under stable environmental conditions is generally supported by empirical studies (Kreiner *et al.*, 2017a). Other factors, like (partial) asexuality, can however reduce the strength of selection against unreduced gametes (see for example Bohutínská *et al.*, 2021). The maintenance of unreduced gamete production in stable environments is likely to be a balance between the rate of mutations favoring their synthesis, and selective forces acting against them.

It appears that the production of unreduced gametes and polyploid individuals and their maintenance in populations is more likely, or even only possible, if polyploidy comes with a fitness advantage. The different theoretical models exploring under which conditions

polyploidy can spread and fix within a diploid population generally support this expectation. The establishment of polyploid individuals within the diploid progenitor population is more likely when polyploid individuals benefit from a higher fitness compared to their diploid progenitors, both in stable conditions (Felber, 1991), or during environmental changes (Oswald & Nuismer, 2011; Yao *et al.*, 2019; Van Drunen & Friedman, 2021), and when other factors limit gene flow among cytotypes, like assortative mating, self-fertilization or clonality (Oswald & Nuismer, 2011; Griswold, 2021; Van Drunen & Friedman, 2021) or restricted pollen and seed dispersals (Li *et al.*, 2004; Baack, 2005).

Theoretical models in which polyploidy can fix neutrally or with an initial decrease in fitness received less attention (Felber, 1991; Husband, 2004; Li *et al.*, 2004; Baack, 2005; Meyers & Levin, 2006; Oswald & Nuismer, 2011). In these conditions, polyploidy can spread or fix only when the population size is small and stochasticity is high (Li *et al.*, 2004), when there is pre-zygotic isolation between cytotypes (Rausch & Morgan, 2005; Oswald & Nuismer, 2011) or when there is a niche differentiation among cytotypes (Rodriguez, 1996).

Surprisingly, the above-mentioned models make the strong assumption that the production of unreduced gametes in the population is a high (much higher than what is found in natural populations, Kreiner *et al.*, 2017a; b) and fixed quantity, that does not vary through time (Felber, 1991; Husband, 2004; Oswald & Nuismer, 2011; Van Drunen & Friedman, 2021). It is known that the production of unreduced gametes is determined genetically (Parrott & Smith, 1986; Tavoletti *et al.*, 1991), varies among individuals within and between population scales (Kreiner *et al.*, 2017a), and is likely to be polygenic (Brownfield & Köhler, 2011; De Storme & Geelen, 2013). It has been pointed out that we are currently lacking clear theoretical and empirical frameworks for the effect of the genetic architecture of unreduced gamete production on the spread and fixation probability of polyploidy (Kreiner *et al.*, 2017a).

In this paper, we investigated theoretically the effect of the genetic architecture of unreduced gametes production in the spread and fixation of polyploidy, when tetraploid individuals are more or less fit than their diploid progenitors. We first built a simple analytical model which allowed us to predict in which conditions the production of unreduced gametes and the formation of tetraploid individuals can evolve by genetic drift. We then built an individual-based simulation model in which the production of unreduced gametes is a quantitative trait that evolves through time by mutation, selection and genetic drift, to test our predictions, and to study the effect of the genetic architecture of unreduced gametes production in the fixation of polyploidy when tetraploids are more or less fit than their diploid progenitors. Overall, we found that when polyploid individuals are less or as fit as their diploid progenitors, the fixation of polyploidy is only possible when genetic drift is stronger than natural selection. Surprisingly, this result remains true even when polyploidy confers a selective advantage, with genetic drift being necessary for tetraploidy to be established, except when polyploids are considerably more fit than their diploid progenitors. Finally, we found that self-fertilization is less beneficial than previously thought, as it increases the selection against unreduced gamete production compared to outcrossing populations when polyploids harbour an initial decrease in fitness.

MATERIAL & METHODS

Verbal prediction

We consider a diploid population in which the proportion of unreduced gametes can evolve through time. The genetic architecture of the trait "proportion of unreduced gametes"

is considered to be the same in male and in female functions. It is hard to judge if the above-mentioned assumption is realistic or not. Male and female functions can slightly vary in their average production of unreduced gametes, but the among-individual variation is high and overlapping in both sexes (Kreiner *et al.*, 2017a).

Consider that the number of unreduced gametes is coded by L bi-allelic loci. The allele A being the ancestral allele and coding for no unreduced gametes, and the allele M at the ith locus increasing the production of unreduced gametes by a value $h.x_i$ and x_i respectively in its heterozygous and homozygous states. In this model, h is the coefficient of dominance of the M allele, and is fixed for all loci. The multi-locus probability of producing an unreduced gamete, p_{UG} is inferred as the sum of genotypic values over the L loci, and depends on the number of mutations carried by an individual (see eq. 1 for details). We assume that the environment effects on unreduced gamete production are null, on average (Falconer & Mackay, 1996). First, we will infer $s_{0,i}$, the coefficient of selection against the allele M at the ith locus at $t = 0$, when the population is uniform and produces no unreduced gametes. We assume that an individual produces γ_{tot} gametes in total. The wildtype homozygote genotype $G_{AA,i}$ (homozygote for the allele A at the ith locus) only produces haploid gametes, such that its absolute fitness is equal to γ_{tot}. To infer $s_{0,i}$, we will consider $G_{AA,i}$ as our reference, and fix its relative fitness to $\omega_{AA} = 1$, while the relative fitness of the homozygote genotype $G_{MM,i}$ (homozygote for the allele M at the ith locus) is equal to $\omega_{MM} = 1 - s_{0,i}$. The number of gametes of $G_{MM,i}$ can be divided as the proportion of haploid gametes γ_n it produces at frequency $p_n = 1\text{-}x_i$, with $\gamma_n = p_n.\gamma_{tot}$, and the proportion of unreduced gametes γ_{2n} produced at frequency $p_{2n} = x_i$, such that $\gamma_{2n} = p_{2n}.\gamma_{tot}$.

Assuming that, initially, the probability of forming a tetraploid offspring is so low due to minority cytotype exclusion, unreduced gametes will not contribute to fitness, so that the relative fitness of $G_{MM,i}$ is easily seen to be $1 - x_i$. It appears that $s_{0,i} = x_i$, and that the smaller

177 the effect of the modifier of unreduced gametes production, the less it will be counter-
178 selected. The average coefficient of selection against unreduced gamete production ($\bar{s_0}$) is the
179 mean of the selection coefficients among loci ($\bar{s_0} = \frac{1}{L}\sum_i^L x_i = \bar{x_i}$). Now we have a way to
180 quantify the strength of selection, we can deduce in which conditions the evolution of
181 unreduced gametes production will be influenced by genetic drift. If we consider that the
182 dynamic of the modifier is mainly driven by genetic drift when N_E. $\bar{s_0} < 1$, then N_E has to be
183 strictly smaller than $1/\bar{s_0}$, where N_E is the effective population size.

184 When selection is strong compared to drift ($N_E > 1/\bar{s_0}$), the frequency of the allele *M* is
185 expected to remain low at all loci, such that the probability of generating tetraploid
186 individuals remains low, and the population is expected to stay diploid. When genetic drift is
187 stronger than selection ($N_E < 1/\bar{s_0}$), the allele *M* can reach higher frequencies, or can even be
188 fixed by chance, such that the probability of generating tetraploid individuals becomes high,
189 and tetraploid individuals have a chance to replace their diploid progenitors. In addition, it has
190 to be noted that the coefficient of selection against unreduced gametes is not fixed, and will
191 evolve through time. Notably, the selection pressure will decrease when the frequency of
192 unreduced gametes (from diploids and tetraploids) will increase, and it can eventually become
193 negative (and then the production of unreduced gametes will be positively selected) when
194 unreduced gametes become dominant in a population. Then, $\bar{s_0}$, and the associated N_E for
195 which the trait is expected to mainly evolve by drift, are the critical values that initiate the
196 system. Any factor decreasing the effective population size, like self-fertilization, is expected
197 to facilitate the increase in the frequency of the deleterious alleles *M* by genetic drift. Self-
198 fertilization (or any other form of assortative mating) also promotes polyploidization by
199 helping polyploids escape the minority cytotype exclusion.

200

Simulation model

In order to test our analytical prediction about the effect of drift in the formation of tetraploid populations, we wrote an individual-based simulation model. We simulated populations with a constant population size of N individuals (N contains both D diploid and T tetraploid individuals), and a constant selfing rate σ. The trait "unreduced gametes production" is coded by L freely recombining bi-allelic loci, with the allele A coding for no unreduced gametes, and the allele M increasing the production of unreduced gametes at the ith locus by a fraction $x_i = 1/L$ in its homozygous state, and by $hx_i = h/L$ in its heterozygous state, h being the coefficient of dominance of the allele M, and being fixed among loci. With this simplification, if an individual carries respectively n and m mutations respectively at heterozygous and homozygous states, the probability of producing an unreduced gamete, p_{UG}, become

$$p_{UG} = \frac{nh+m}{L} \text{ , (eq. 1)}$$

and varied between 0 (the individual is homozygote for the allele A at all loci) and 1 (the individual is homozygote for the allele M at all loci). As x_i is set to $1/L$ at all loci, $\overline{s_0} = 1/L$, and genetic drift is stronger than selection when $N_E < L$. For each individual and at each generation, an environmental effect is drawn in a Gaussian distribution of mean 0 and variance V_E of 0.005, in order to introduce variation in unreduced gametes production. The fitness of diploid individuals is fixed to $W_{2x} = 1$, while the fitness of neo-tetraploid individuals is fixed within a simulation, but can vary from $W_{4x} = 0.75$ to 1.25 among simulations, to fit with empirical data demonstrating initial fitness decrease in neo-polyploids due to their problems in cellular management and meiosis (Porturas *et al.*, 2019; Clo & Kolář, 2021), or higher fitness in stressful conditions (Chao *et al.*, 2013; Ruiz *et al.*, 2016). We assume random

mating opportunities among individuals of any ploidy. Only diploid and tetraploid cytotypes are viable, triploids and higher polyploids are assumed to be not viable (Ramsey & Schemske, 1998; Burton & Husband, 2000; Husband & Sabara, 2003). Note that this hypothesis is conservative, as viable triploid individuals are expected to favour the establishment of tetraploid populations (Husband, 2004). However, a (virtual) lack of triploids in many mixed-ploidy systems (Kolář *et al.*, 2017) demonstrates that such a conservative model is realistic for many species. The population evolves during 20K generations, and the frequencies of the allele *M*, of unreduced gametes and of tetraploid individuals are followed.

We initiated each simulation from a population of genetically identical diploid individuals that are homozygous for the allele *A* at all loci (p_{UG} = 0 at t = 0). The life cycle can be summarized in five successive events. First, there is a fitness-dependent choice of the first parent (selection), followed by mating-type choice (selfing versus outcrossing at rates σ and [1-σ], respectively). In the case of outcrossing, the choice of the second parent is also fitness-dependent. Selection takes place as follows: an individual is sampled randomly among the *N* individuals, but its sampling probability is weighted by its fitness. The probabilities of sampling a diploid (p_{2X}) or a tetraploid (p_{4X}) individual are respectively

$$p_{2X} = \frac{D.W_{2x}}{D.W_{2x}+T.W_{4x}}, \text{ (eq 2.1)}$$

and

$$p_{4X} = \frac{T.W_{4x}}{D.W_{2x}+T.W_{4x}}. \text{ (eq 2.2)}$$

With these equations, it appears that tetraploids are as likely to be sampled as diploids when they represent respectively 57%, 50% or 44% of the population when W_{4x} is respectively equal to 0.75, 1 or 1.25. For mixed mating regime, a number is sampled in a uniform law comprised between 0 and 1, if the number is smaller (respectively higher) than 0.5, the

selected individual reproduced by outcrossing (respectively by selfing). Once the two parents are chosen, the type of gametes (reduced or unreduced) they will produce is chosen. For each reproducer, one number is sampled in a uniform distribution between 0 and 1. If p_{UG} is higher than the sampled value, the reproducer generates an unreduced gamete (note that a tetraploid individual can also produce unreduced gametes, see below for details). For reduced (n) gametes (n = x and n = 2x gametes for diploid and tetraploid individuals, respectively), one allele per locus of each sister chromatid is sampled randomly. This phase is then followed by the introduction of mutations from allele *A* to allele *M* and *vice versa*, the number of which is sampled from a Poisson distribution with parameter U (with $U = \mu L$, where μ is the per-locus mutation rate and L is the number of loci underlying the quantitative trait). If the resulting offspring is not diploid or tetraploid, it will not contribute to the next generation. The reproduction phase stops when N viable offspring are formed.

Parameter sets tested

The number of loci underlying the production of unreduced gametes can vary from L = 2 to 200. The population size can be equal to N = 50, 100 or 200. The associated effective population size is equal to $N_E = N(2-\sigma)/2$ (Glémin & Ronfort, 2013). Bigger population sizes (*e.g.* $N > 2000$) do not change the results qualitatively, but the simulation time becomes limiting, and the number of loci needed for the trait to evolve by genetic drift becomes unrealistically high. We tested three mutation rates, with the diploid genomic mutation rate (U) being equal to 0.5, 0.05 or 0.005, which is the range of what is found in natural populations (Halligan & Keightley, 2009). For simplicity, the coefficient of dominance of the alleles *M* can be equal to 0 (fully recessive: none of the heterozygote genotypes in diploids or

tetraploids produce unreduced gametes), 0.5 (co-dominant, the diploid and tetraploid heterozygotes genotypes at the *i*th locus are: $G_{AM,i}$, $G_{AAAM,i}$, $G_{AAMM,i}$ and $G_{AMMM,i}$ produce unreduced gametes at frequencies 0.5/*L*, 0.25/*L*, 0.5/*L* and 0.75/*L*, respectively) or 1 (fully dominant: any heterozygote genotypes in diploids or tetraploids produce unreduced gametes at a frequency 1/*L* at a given locus). Finally, we tested three selfing rates with σ being equal to 0, 0.5 or 1. The selfing rate is the same in diploids and tetraploids. As mentioned before, tetraploid individuals can be more or less fit compared to their diploid progenitors, reflecting the fitness (dis)advantage associated with tetraploidy in standard *vs* disturbed environments. For each parameter, we performed $n = 100$ replicates. A summary of the parameters tested and their values can be found in table 1.

Table 1. Description of the model parameters, their abbreviations, and tested values in simulations.

Parameter	Abbreviation	Value(s)
Population size	N	50, 100, 200
Number of loci	L	From 2 to 200
Diploid genomic mutation rate	U	0.005, 0.05 or 0.5
Fitness of diploid individuals	W_{2x}	1
Fitness of tetraploid individuals	W_{4x}	0.75, 1, 1.25 or 2
Selfing rate	σ	0, 0.5 or 1
Probability of sampling an unreduced gamete, a diploid or tetraploid individual	p_{UG}, p_{2x}, p_{4x}	From 0 to 1
Coefficient of selection at $t = 0$	$\overline{s_0}$	From 0.5 to 0.005
Dominance's coefficient of allele M	h	0, 0.5, 1
V_E of environmental effects	V_E	0.005

RESULTS

The effects of fitness and genetics on the probability of fixation of polyploidy

As predicted analytically, when considering tetraploids being less fit than diploids the fixation of polyploidy is only possible when genetic drift is stronger than natural selection against the unreduced gamete formation (Figure 1, left panel). This finding remains true, even when tetraploid individuals are as fit as their diploid progenitors (Figure 1, middle panel). If tetraploids are slightly fitter than their diploid progenitors, fixation of the tetraploidy occurs either when drift is stronger or as strong as selection, but never when selection is stronger than drift (Figure 1, right panel). Furthermore, when tetraploid individuals are as fit as or slightly fitter than diploids, the fixation of polyploids is faster compared to the case in which tetraploids are less fit than diploids (Figure 1 middle and right panels compared to the left panel). Polyploidy can fix without drift only when it comes with an important fitness advantage ($W_{tetra} = 2W_{diplo}$, Figure 2). More dominant mutations and higher mutation rates increased the production of unreduced gametes and then the probability of fixing polyploidization (Figure S1 & S2).

The production of unreduced gametes within populations

The production of unreduced gametes in diploid individuals strongly depends on two parameters, the mutation rate for unreduced gamete production, and the strength of drift compare to selection (Figure 3). As expected, higher mutation rates lead to higher production

of unreduced gametes within populations (Figure 3, U = 0.005 *vs.* U = 0.05). For example, when selection is stronger than drift, the production of unreduced gametes in diploid individuals is on average 2.8% when U = 0.005, and is on average 17.1% when U = 0.05. As predicted, the higher the level of genetic drift, the higher the number of unreduced gametes produced by individuals (Figure 3). For example, for U = 0.05, the production of unreduced gametes in diploids is on average 17.1% when selection is stronger than drift, and around 30.4% when drift is stronger than selection. The same pattern is observable for smaller mutation rates (Figure 3). The fitness of tetraploid individuals slightly affects the production of unreduced gametes in diploids, the fitter the tetraploids, the higher the production of unreduced gametes in diploids (Figure S3). The production of unreduced gametes remains comparable in diploids and tetraploid individuals, on average (Figure S4).

Effect of the mating system

Self-fertilization appears to increase the strength of selection against unreduced gamete formation when polyploidy is associated with a fitness cost, despite its positive effect on the strength of genetic drift. When genetic drift is stronger than selection and polyploidy has a fitness cost, only obligate outcrossing populations are able to fix polyploidy (Figure 4 left panel). When polyploidy comes with a fitness advantage over diploidy, we confirmed the classical expectation that self-fertilization increases the probability of fixing polyploidy, due to the escaping from the minority cytotype exclusion, when genetic drift is stronger (Figure 4 middle panel) or weaker (Figure 4 right panel) than natural selection.

DISCUSSION

Genetic drift is a major component of the probability of fixation of polyploidy

One major result of our theoretical model is that genetic drift is important for the evolution of polyploidy, even when autotetraploids are fitter than their diploid progenitors. This result indicates that the initial frequency-dependent selection is the strongest barrier to polyploids establishment.

Our predictions offer a new and alternative view on the establishment of polyploidy in populations. Even if polyploid lineages are more likely to survive given their potential fitness advantage over their diploid progenitors (Van de Peer *et al.*, 2021), this does not necessarily mean such a slight adaptive advantage has been the cause for their establishment. Harsher environment conditions and/or periods of drastic environmental changes, extreme latitudes, and range margins are also expected to decrease the demographic and effective population sizes of species (Brown, 1984), offering the possibility for polyploidization to fix by genetic drift. It can explain why polyploid species are contemporarily more frequently found at range margins (Levin, 1975; Fowler & Levin, 1984; Felber, 1991; Rojas-Andrés *et al.*, 2020), extreme environments (Brochmann *et al.*, 2004), extreme latitudes (Rice *et al.*, 2019) and why they are frequently invasive (Te Beest *et al.*, 2012). It can also help to understand why polyploidy can sometimes be associated with higher extinction rates compared to diploidy (Mayrose *et al.*, 2011, 2015; Arrigo & Barker, 2012), if polyploidy has been initially fixed despite an initial decrease in fitness. It is also important to note that drift and natural selection are not exclusive explanations, and both forces may increase the probability of establishment of polyploid lineages under certain scenarios. Genetic drift can help in increasing the frequency of genes favouring unreduced gametes production and escaping minority cytotype

exclusion, while selection will then maintain them at high frequency if polyploid individuals are more fit than their diploid progenitors.

Direct evidence of the role of drift in the evolution of polyploidy is however hard to find and could be hard to produce in the future. Nevertheless, and as a support of our model, the few experiments testing for local adaptation of polyploid lineages led to mitigated results, with either no (Baack & Stanton, 2005; Buggs & Pannell, 2007; Martin & Husband, 2013; McIntyre & Strauss, 2017) or a weak (Raabová *et al.*, 2008; Duchoslav *et al.*, 2017) fitness advantages of polyploids in their local sites. Note that testing for local adaptation patterns can be a weak way of testing for the effect of drift or selection in polyploidy establishment, because a fitness advantage is not always sufficient for polyploidy to fix, and because polyploidization can fix with an initial decrease in fitness with the help of genetic drift, and adaptation can occur later after establishment.

Comparing the number of unreduced gametes with empirical results

In our simulations, the average number of unreduced gametes within diploid populations that tetraploids are unable to establish (*i.e.* population in which selection is stronger than drift) strongly depends on the mutation rate for unreduced gametes production. Empirical estimates made in diploid populations range from 0.2% to 11.6% (Bretagnolle, 2001; Ramsey, 2007; Kron & Husband, 2009; Sheidai *et al.*, 2009; Kovalsky & Solís Neffa, 2012, 2016; Herben *et al.*, 2016; Kreiner *et al.*, 2017b), with an average number of 2.6% (Kreiner *et al.*, 2017a). In our simulations, we are obtaining similar values only when the mutation rate for unreduced gametes is low (Figure 3 & S3) and when selection is stronger

than drift, suggesting that it is also probably the case in the few natural populations that have been studied.

However, the number of unreduced gametes needed for polyploidy to fix in an initially diploid population is much higher in our simulations (at least 15%) than what is found in natural diploid populations (Kreiner *et al.*, 2017a). Our results suggest that, even if polyploid individuals are fitter than their diploid progenitors, the number of unreduced gametes found in natural diploid populations would theoretically not allow polyploid individuals to invade. This result challenges the idea that the production of unreduced gamete in diploid populations may be selected to enhance polyploid production in stressful conditions or in habitats where they are most expected to thrive (Stebbins, 1950; Vanneste *et al.*, 2014b; Mason & Pires, 2015). The production of unreduced gametes remains similar in diploid and tetraploid individuals after the fixation of polyploidy. This is mainly because of the assumptions we made in our model, notably that after fixation of tetraploidy, there are no other cytotypes possible, making intermediate rates of unreduced gamete production not costly in tetraploid populations, in our simulation model.

Pre-mating isolation among cytotypes and polyploidy establishment

We found that pre-mating isolation among cytotypes promotes polyploidization only when tetraploid individuals are as fit or fitter than their diploid progenitors. This stems from the fact that selfing lineages that tend to produce more unreduced gametes will also produce more tetraploid offspring with a lower fitness, compared to other selfing lineages that produce more frequently diploid offspring with a higher fitness. As neo-polyploid lineages are generally less fit than their diploid progenitors (Porturas *et al.*, 2019; Clo & Kolář, 2021), it

could be expected that the evolution of self-fertilization particularly, and other forms of assortative mating in general, should be limited in empirical data. As a confirmation, Husband *et al.* (2008) found that autopolyploid lineages generally have smaller selfing rates than diploid species. Other forms of assortative mating, such as delayed phenology among cytotypes, are also rarely observed at a large phylogenetic scale (Porturas *et al.*, 2019; Clo & Kolář, 2021). Our results also fit well with empirical data that did not find an effect of the mating system on the frequency of unreduced gametes in natural populations (Kreiner *et al.*, 2017b).

Considerations related to the genetic architecture of unreduced gametes production

It is first important to notice that little is known about the genetic architecture of the production of unreduced gametes in natural populations (Brownfield & Köhler, 2011; De Storme & Geelen, 2013). As a consequence, we cannot precisely quantify for which effective population sizes the production of unreduced gametes can occur by genetic drift when polyploidization is detrimental to individuals' fitness. It is known, however, that there are several pre-meiotic (Fernández & Neffa, 2004; Kim *et al.*, 2009; Mason *et al.*, 2011), meiotic (Bretagnolle & Thompson, 1995; Pécrix *et al.*, 2011; De Storme *et al.*, 2012) and post-meiotic (Bastiaanssen *et al.*, 1998; Ramanna & Jacobsen, 2003) pathways that lead to unreduced gametes production. For genetic drift to have a chance to play a role in polyploidization events, it is necessary that these physiological pathways are not all governed by a few major effect mutations (Fig. 2). If there are enough loci with small effects, then it becomes more likely that genetic drift can play a role in the production of unreduced gametes and polyploidization events, independently of the (mal)adaptive value of polyploid individuals.

Limitations of the model

Although we tried to select realistic parameter settings (e.g. for the fitness of polyploids compared to diploids, the mutation rate of the trait), due to limited empirical proofs, some assumptions could be unrealistic. This could make our predictions and further empirical research unreliable. The first limitation is that we considered that the genetic architecture of unreduced gametes production is the same for the male and female functions. While the few empirical data available do not reject this assumption (Kreiner *et al.*, 2017a), it may not be true in some species (De Storme & Geelen, 2013). Our predictions would become unreliable if the production of unreduced gametes is coded by a few numbers of major effect loci in one sex, and by a large number of small effect loci in the other. In such a case, polyploidization is unlikely to evolve by genetic drift, as it would be constrained by one of the two sexes. Another simplification is that we hypothesized that triploid and higher polyploid levels are not viable. While the hypothesis is justified empirically for multiple plant species (Ramsey & Schemske, 1998; Burton & Husband, 2000; Husband & Sabara, 2003; Kolář *et al.*, 2017), there are also species with frequent natural triploidy (Trávníček *et al.*, 2011; Čertner *et al.*, 2019), and it has been shown theoretically that viable triploids favour the establishment of polyploid populations (Husband, 2004). Relaxing our hypothesis would not change our conclusions qualitatively, since it would decrease the selection pressure against the production of unreduced gametes, and favour the evolution of polyploidy by drift for a wider set of parameters. Finally, we considered a fixed population size in our simulations, but the evolution toward polyploidy can affect the populations' survival in different ways. First, if the fitness of polyploids is smaller than the one of diploids, the mean population fitness will decrease, and the growth rate can become smaller than one, leading to the population's

decline and extinction. Second, genetic drift can also lead to the fixation of deleterious mutations, if the average coefficient of selection of deleterious mutations is similar or smaller to the one of mutations involved in the production of unreduced gametes. In such a case, diploids and/or tetraploids can accumulate deleterious mutations, potentially leading to the extinction of populations by mutational meltdown (Lynch *et al.*, 1995).

CONCLUSIONS

In this paper, we investigated theoretically the effect of genetic drift on the probability of fixation of polyploidy in a population under a scenario assuming selection against unreduced gamete production, a generally maladaptive trait. The main conclusions of our model are that, for polyploidization to fix, genetic drift is necessary, independently of the selective (dis)advantage of tetraploid individuals over their diploid progenitors. Thus, our simulation models suggest that the cost of being the minority cytotype in the population of origin is stronger than a slight increase in fitness associated with polyploidy. The effective population size necessary for the production of unreduced gametes to evolve by genetic drift depends on the average effects of the loci coding for the trait. This implies that if unreduced gamete production is coded by a few major effect loci, or if a species has a large effective population size, polyploidization is unlikely to occur without a strong selective advantage. Selfing rate, and probably other forms of assortative mating, seems to be deterministic in the fixation of polyploidy only when polyploids do not suffer from an initial decrease in fitness. However, we cannot discard that other reproductive interactions, such as fertile intermediate hybrids, can affect the outcome. On the opposite, any factors reducing the demographic and effective population size are expected to favour polyploidization, independently of the

selective (dis)advantages of tetraploids compared to diploids. Our model thus offers new and non-exclusive explanations of why polyploidization is associated with periods of extinction and environmental turmoil, and harsh and/or disturbed environments (Rice *et al.*, 2019; Van de Peer *et al.*, 2021).

Testing our predictions empirically will remain however quite challenging. Although we know that the formation of unreduced gametes is determined genetically and that it exhibits heritable variance (Parrott & Smith, 1986; Tavoletti *et al.*, 1991), little is known about the number of genes underlying the trait (Brownfield & Köhler, 2011; De Storme & Geelen, 2013; Kreiner *et al.*, 2017a). In order to test our hypothesis, one of the next steps would be to investigate at a large taxonomic scale the genetic architecture of unreduced gamete production. The other challenge is that, when polyploidization occurs, the fixation time of polyploidy is really small, increasing the probability of missing the evolution of polyploidy. Invasive species could be interesting model species, as founding events are expected to decrease the effective population size of populations, and it is more likely for polyploidy to evolve during the invasion process. Simón-Porcar *et al.* (2017) recently found a mixed-ploidy population of the invasive diploid *Mimulus guttatus* in the UK that could be a good system for following our investigation of the role of drift in polyploidy evolution.

BIBLIOGRAPHY

Arrigo, N. & Barker, M.S. 2012. Rarely successful polyploids and their legacy in plant genomes. *Current opinion in plant biology* **15**: 140–146. Elsevier.

Baack, E.J. 2005. To succeed globally, disperse locally: effects of local pollen and seed dispersal on tetraploid establishment. *Heredity* **94**: 538–546. Nature Publishing Group.

519 Baack, E.J. & Stanton, M.L. 2005. Ecological factors influencing tetraploid speciation in snow
520 buttercups (Ranunculus adoneus): niche differentiation and tetraploid establishment.
521 *Evolution* **59**: 1936–1944. Wiley Online Library.

522 Barker, M.S., Arrigo, N., Baniaga, A.E., Li, Z. & Levin, D.A. 2016. On the relative abundance of
523 autopolyploids and allopolyploids. *New Phytologist* **210**: 391–398. Wiley Online Library.

524 BASTIAANSSEN†, VAN DEN BERG, LINDHOUT, JACOBSEN, & RAMANNA*. 1998. Postmeiotic
525 restitution in 2n-egg formation of diploid potato. *Heredity* **81**: 20–27. John Wiley & Sons, Ltd.

526 Bohutínská, M., Alston, M., Monnahan, P., Mandáková, T., Bray, S., Paajanen, P., *et al.* 2021. Novelty
527 and convergence in adaptation to whole genome duplication. *Molecular biology and*
528 *evolution* **38**: 3910–3924. Oxford University Press.

529 Bomblies, K. 2020. When everything changes at once: finding a new normal after genome
530 duplication. *Proceedings of the Royal Society B* **287**: 20202154. The Royal Society.

531 Bretagnolle, F. 2001. Pollen production and spontaneous polyploidization in diploid populations of
532 Anthoxanthum alpinum. *Biological Journal of the Linnean Society* **72**: 241–247. Oxford
533 University Press.

534 Bretagnolle, F. & and THOMPSON, J.D. 1995. Gametes with the somatic chromosome number:
535 mechanisms of their formation and role in the evolution of autopolyploid plants. *New*
536 *Phytologist* **129**: 1–22. Wiley Online Library.

537 Brochmann, C., Brysting, A.K., Alsos, I.G., Borgen, L., Grundt, H.H., Scheen, A.-C., *et al.* 2004.
538 Polyploidy in arctic plants. *Biological journal of the Linnean society* **82**: 521–536. Oxford
539 University Press.

540 Brown, J.H. 1984. On the relationship between abundance and distribution of species. *The american*
541 *naturalist* **124**: 255–279. University of Chicago Press.

542 Brownfield, L. & Köhler, C. 2011. Unreduced gamete formation in plants: mechanisms and prospects.
543 *Journal of experimental botany* **62**: 1659–1668. Oxford University Press.

544 Buggs, R.J. & Pannell, J.R. 2007. Ecological differentiation and diploid superiority across a moving
545 ploidy contact zone. *Evolution* **61**: 125–140. Wiley Online Library.

546 Burton, T.L. & Husband, B. 2000. Fitness differences among diploids, tetraploids, and their triploid
547 progeny in Chamerion angustifolium: mechanisms of inviability and implications for polyploid
548 evolution. *Evolution* **54**: 1182–1191. Wiley Online Library.

549 Čertner, M., Kúr, P., Kolář, F. & Suda, J. 2019. Climatic conditions and human activities shape diploid–
550 tetraploid coexistence at different spatial scales in the common weed Tripleurospermum
551 inodorum (Asteraceae). *Journal of Biogeography* **46**: 1355–1366. Wiley Online Library.

552 Chao, D.-Y., Dilkes, B., Luo, H., Douglas, A., Yakubova, E., Lahner, B., *et al.* 2013. Polyploids exhibit
553 higher potassium uptake and salinity tolerance in Arabidopsis. *Science* **341**: 658–659.
554 American Association for the Advancement of Science.

555 Clo, J. & Kolář, F. 2021. Short- and long-term consequences of genome doubling: a meta-analysis.
556 *American Journal of Botany* **108**: 2315–2322. John Wiley & Sons, Ltd.

557 Comai, L. 2005. The advantages and disadvantages of being polyploid. *Nature Reviews Genetics* **6**:
558 836–846.

559 De Storme, N., Copenhaver, G.P. & Geelen, D. 2012. Production of diploid male gametes in
560 Arabidopsis by cold-induced destabilization of postmeiotic radial microtubule arrays. *Plant*
561 *physiology* **160**: 1808–1826. American Society of Plant Biologists.

562 De Storme, N. & Geelen, D. 2013. Sexual polyploidization in plants – cytological mechanisms and
563 molecular regulation. *New Phytologist* **198**: 670–684. John Wiley & Sons, Ltd.

564 Doyle, J.J. & Coate, J.E. 2019. Polyploidy, the nucleotype, and novelty: the impact of genome
565 doubling on the biology of the cell. *International Journal of Plant Sciences* **180**: 1–52.
566 University of Chicago Press Chicago, IL.

567 Duchoslav, M., Fialová, M. & Jandová, M. 2017. The ecological performance of tetra-, penta-and
568 hexaploid geophyte Allium oleraceum in reciprocal transplant experiment may explain the
569 occurrence of multiple-cytotype populations. *Journal of Plant Ecology* **10**: 569–580. Oxford
570 University Press.

571 Falconer, D.S. & Mackay, T.F.C. 1996. Introduction to quantitative genetics. Essex. *UK: Longman*
572 *Group*.

573 Felber, F. 1991. Establishment of a tetraploid cytotype in a diploid population: effect of relative
574 fitness of the cytotypes. *Journal of evolutionary biology* **4**: 195–207. Wiley Online Library.

575 Fernández, A. & Neffa, V.G.S. 2004. Genomic relationships between Turnera krapovickasii (2x, 4x)
576 and T. ulmifolia (6x)(Turneraceae, Turnera). *Caryologia* **57**: 45–51. Taylor & Francis.

577 Fowler, N.L. & Levin, D.A. 1984. Ecological constraints on the establishment of a novel polyploid in
578 competition with its diploid progenitor. *The American Naturalist* **124**: 703–711. University of
579 Chicago Press.

580 Glémin, S. & Ronfort, J. 2013. Adaptation and maladaptation in selfing and outcrossing species: new
581 mutations versus standing variation. *Evolution: International Journal of Organic Evolution* **67**:
582 225–240. Wiley Online Library.

583 Gregory, T.R. & Mable, B.K. 2005. Polyploidy in animals. In: *The evolution of the genome*, pp. 427–
584 517. Elsevier.

585 Griswold, C.K. 2021. The effects of migration load, selfing, inbreeding depression, and the genetics of
586 adaptation on autotetraploid versus diploid establishment in peripheral habitats. *Evolution*
587 **75**: 39–55. Wiley Online Library.

588 Halligan, D.L. & Keightley, P.D. 2009. Spontaneous mutation accumulation studies in evolutionary
589 genetics. *Annual Review of Ecology, Evolution, and Systematics* **40**: 151–172. Annual Reviews.

590 Hannweg, K., Steyn, W. & Bertling, I. 2016. In vitro-induced tetraploids of Plectranthus esculentus are
591 nematode-tolerant and have enhanced nutritional value. *Euphytica* **207**: 343–351. Springer.

592 Harlan, J.R. & deWet, J.M. 1975. On Ö. Winge and a prayer: the origins of polyploidy. *The botanical*
593 *review* **41**: 361–390. Springer.

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

Hias, N., Svara, A. & Keulemans, J.W. 2018. Effect of polyploidisation on the response of apple (Malus$\times$ domestica Borkh.) to Venturia inaequalis infection. *European Journal of Plant Pathology* **151**: 515–526. Springer.

Husband, B.C. 2004. The role of triploid hybrids in the evolutionary dynamics of mixed-ploidy populations. *Biological Journal of the Linnean Society* **82**: 537–546.

Husband, B.C., Ozimec, B., Martin, S.L. & Pollock, L. 2008. Mating consequences of polyploid evolution in flowering plants: current trends and insights from synthetic polyploids. *International Journal of Plant Sciences* **169**: 195–206. The University of Chicago Press.

Husband, B.C. & Sabara, H.A. 2003. Reproductive isolation and their diploid autotetraploids Chamerion progenitors in fireweed, Chamerion angustifolium (Onagraceae). *New Phytologist* **161**: 703–713.

Kim, J.S., Oginuma, K. & Tobe, H. 2009. Syncyte formation in the microsporangium of Chrysanthemum (Asteraceae): a pathway to infraspecific polyploidy. *Journal of Plant Research* **122**: 439–444. Springer.

Kolář, F., Čertner, M., Suda, J., Schönswetter, P. & Husband, B.C. 2017. Mixed-ploidy species: progress and opportunities in polyploid research. *Trends in Plant Science* **22**: 1041–1055. Elsevier.

Kovalsky, I.E. & Solís Neffa, V.G. 2012. Evidence of 2n microspore production in a natural diploid population of Turnera sidoides subsp. carnea and its relevance in the evolution of the T. sidoides (Turneraceae) autopolyploid complex. *Journal of plant research* **125**: 725–734. Springer.

Kovalsky, I.E. & Solís Neffa, V.G. 2016. Evidence of the production of 2n eggs in diploid plants of the autopolyploid complex Turnera sidoides L.(Passifloraceae). *Plant systematics and evolution* **302**: 357–366. Springer.

Kreiner, J.M., Kron, P. & Husband, B.C. 2017a. Evolutionary dynamics of unreduced gametes. *Trends in Genetics* **33**: 583–593. Elsevier.

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

Kron, P. & Husband, B.C. 2009. Hybridization and the reproductive pathways mediating gene flow between native Malus coronaria and domestic apple, M. domestica. *Botany* **87**: 864–874.

Levin, D.A. 1975. Minority cytotype exclusion in local plant populations. *Taxon* **24**: 35–43. Wiley Online Library.

Li, B.-H., Xu, X.-M. & Ridout, M.S. 2004. Modelling the establishment and spread of autotetraploid plants in a spatially heterogeneous environment. *Journal of Evolutionary Biology* **17**: 562–573. Wiley Online Library.

633 Lynch, M., Conery, J. & Burger, R. 1995. Mutation accumulation and the extinction of small
634 populations. *The American Naturalist* **146**: 489–518. University of Chicago Press.

635 Mable, B.K. 2004. 'Why polyploidy is rarer in animals than in plants': myths and mechanisms.
636 *Biological Journal of the Linnean Society* **82**: 453–466. Oxford University Press.

637 Martin, S.L. & Husband, B.C. 2013. Adaptation of diploid and tetraploid Chamerion angustifolium to
638 elevation but not local environment. *Evolution* **67**: 1780–1791. Wiley Online Library.

639 Mason, A.S., Nelson, M.N., Yan, G. & Cowling, W.A. 2011. Production of viable male unreduced
640 gametes in Brassica interspecific hybrids is genotype specific and stimulated by cold
641 temperatures. *BMC plant biology* **11**: 1–13. BioMed Central.

642 Mason, A.S. & Pires, J.C. 2015. Unreduced gametes: meiotic mishap or evolutionary mechanism?
643 *Trends in Genetics* **31**: 5–10. Elsevier.

644 Mayrose, I., Zhan, S.H., Rothfels, C.J., Arrigo, N., Barker, M.S., Rieseberg, L.H., *et al.* 2015. Methods
645 for studying polyploid diversification and the dead end hypothesis: a reply to Soltis et
646 al.(2014). *New Phytologist* **206**: 27–35. JSTOR.

647 Mayrose, I., Zhan, S.H., Rothfels, C.J., Magnuson-Ford, K., Barker, M.S., Rieseberg, L.H., *et al.* 2011.
648 Recently formed polyploid plants diversify at lower rates. *Science* **333**: 1257–1257. American
649 Association for the Advancement of Science.

650 McIntyre, P.J. & Strauss, S. 2017. An experimental test of local adaptation among cytotypes within a
651 polyploid complex. *Evolution* **71**: 1960–1969. Wiley Online Library.

652 Meyers, L.A. & Levin, D.A. 2006. On the abundance of polyploids in flowering plants. *Evolution* **60**:
653 1198–1206. Wiley Online Library.

654 Oberlander, K.C., Dreyer, L.L., Goldblatt, P., Suda, J. & Linder, H.P. 2016. Species-rich and polyploid-
655 poor: Insights into the evolutionary role of whole-genome duplication from the Cape flora
656 biodiversity hotspot. *American Journal of Botany* **103**: 1336–1347. Wiley Online Library.

657 Oswald, B.P. & Nuismer, S.L. 2011. A unified model of autopolyploid establishment and evolution.
658 *The American Naturalist* **178**: 687–700. University of Chicago Press Chicago, IL.

659 Otto, S.P. 2007. The evolutionary consequences of polyploidy. *Cell* **131**: 452–462. Elsevier.

660 Parisod, C., Holderegger, R. & Brochmann, C. 2010. Evolutionary consequences of autopolyploidy.
661 *New phytologist* **186**: 5–17. Wiley Online Library.

662 Parrott, W.A. & Smith, R.R. 1986. Evidence for the existence of endosperm balance number in the
663 true clovers (Trifolium spp.). *Canadian journal of genetics and cytology* **28**: 581–586. NRC
664 Research Press Ottawa, Canada.

665 Pécrix, Y., Rallo, G., Folzer, H., Cigna, M., Gudin, S. & Le Bris, M. 2011. Polyploidization mechanisms:
666 temperature environment can induce diploid gamete formation in Rosa sp. *Journal of*
667 *experimental botany* **62**: 3587–3597. Oxford University Press.

668 Porturas, L.D., Anneberg, T.J., Curé, A.E., Wang, S., Althoff, D.M. & Segraves, K.A. 2019. A meta-
669 analysis of whole genome duplication and the effects on flowering traits in plants. *American*
670 *journal of botany* **106**: 469–476. Wiley Online Library.

671 Raabová, J., Fischer, M. & Münzbergová, Z. 2008. Niche differentiation between diploid and
672 hexaploid Aster amellus. *Oecologia* **158**: 463–472. Springer.

673 Ramanna, M.S. & Jacobsen, E. 2003. Relevance of sexual polyploidization for crop improvement–A
674 review. *Euphytica* **133**: 3–8. Springer.

675 Ramsey, J. 2007. Unreduced gametes and neopolyploids in natural populations of Achillea borealis
676 (Asteraceae). *Heredity* **98**: 143–150. Nature Publishing Group.

677 Ramsey, J. & Schemske, D.W. 2002. Neopolyploidy in flowering plants. *Annual review of ecology and*
678 *systematics* **33**: 589–639. Annual Reviews 4139 El Camino Way, PO Box 10139, Palo Alto, CA
679 94303-0139, USA.

680 Ramsey, J. & Schemske, D.W. 1998. Pathways, mechanisms, and rates of polyploid formation in
681 flowering plants. *Annual review of ecology and systematics* **29**: 467–501. Annual Reviews
682 4139 El Camino Way, PO Box 10139, Palo Alto, CA 94303-0139, USA.

683 Rausch, J.H. & Morgan, M.T. 2005. THE EFFECT OF SELF-FERTILIZATION, INBREEDING DEPRESSION,
684 AND POPULATION SIZE ON AUTOPOLYPLOID ESTABLISHMEN. *Evolution* **59**: 1867–1875. Wiley
685 Online Library.

686 Rice, A., Šmarda, P., Novosolov, M., Drori, M., Glick, L., Sabath, N., *et al.* 2019. The global
687 biogeography of polyploid plants. *Nature Ecology & Evolution* **3**: 265–273. Nature Publishing
688 Group.

689 Rodriguez, D.J. 1996. A model for the establishment of polyploidy in plants. *The American Naturalist*
690 **147**: 33–46. University of Chicago Press.

691 Rojas-Andrés, B.M., Padilla-García, N., de Pedro, M., López-González, N., Delgado, L., Albach, D.C., *et*
692 *al.* 2020. Environmental differences are correlated with the distribution pattern of cytotypes
693 in Veronica subsection Pentasepalae at a broad scale. *Annals of botany* **125**: 471–484. Oxford
694 University Press US.

695 Ruiz, M., Quiñones, A., Martínez-Cuenca, M.-R., Aleza, P., Morillon, R., Navarro, L., *et al.* 2016.
696 Tetraploidy enhances the ability to exclude chloride from leaves in carrizo citrange seedlings.
697 *Journal of plant physiology* **205**: 1–10. Elsevier.

698 Sheidai, M., Jafari, S., Taleban, P. & Keshavarzi, M. 2009. Cytomixis and unreduced pollen grain
699 formation in Alopecurus L. and Catbrosa Beauv.(Poaceae). *Cytologia* **74**: 31–41. Japan
700 Mendel Society, International Society of Cytology.

701 Simón-Porcar, V.I., Silva, J.L., Meeus, S., Higgins, J.D. & Vallejo-Marín, M. 2017. Recent
702 autopolyploidization in a naturalized population of Mimulus guttatus (Phrymaceae).
703 *Botanical Journal of the Linnean Society* **185**: 189–207. Oxford University Press.

704 Soltis, D.E., Soltis, P.S., Schemske, D.W., Hancock, J.F., Thompson, J.N., Husband, B.C., *et al.* 2007.
705 Autopolyploidy in angiosperms: have we grossly underestimated the number of species?
706 *Taxon* **56**: 13–30. Wiley Online Library.

707 Stebbins, G.L. 1971. *Processes of organic evolution*.

708 Stebbins, G.L. 1950. *Variation and evolution in plants*. Columbia University Press.

709 Tavoletti, S., Mariani, A. & Veronesi, F. 1991. Phenotypic recurrent selection for 2n pollen and 2n egg
710 production in diploid alfalfa. *Euphytica* **57**: 97–102. Springer.

711 Tayalé, A. & Parisod, C. 2013. Natural pathways to polyploidy in plants and consequences for genome
712 reorganization. *Cytogenetic and Genome Research* **140**: 79–96. Karger Publishers.

713 Te Beest, M., Le Roux, J.J., Richardson, D.M., Brysting, A.K., Suda, J., Kubešová, M., *et al.* 2012. The
714 more the better? The role of polyploidy in facilitating plant invasions. *Annals of botany* **109**:
715 19–45. Oxford University Press.

716 Trávníček, P., Dočkalová, Z., Rosenbaumová, R., Kubátová, B., Szeląg, Z. & Chrtek, J. 2011. Bridging
717 global and microregional scales: ploidy distribution in Pilosella echioides (Asteraceae) in
718 central Europe. *Annals of Botany* **107**: 443–454. Oxford University Press.

719 Van de Peer, Y., Ashman, T.-L., Soltis, P.S. & Soltis, D.E. 2021. Polyploidy: an evolutionary and
720 ecological force in stressful times. *The Plant Cell* **33**: 11–26. Oxford University Press.

721 Van de Peer, Y., Mizrachi, E. & Marchal, K. 2017. The evolutionary significance of polyploidy. *Nature*
722 *Reviews Genetics* **18**: 411. Nature Publishing Group.

723 Van Drunen, W.E. & Friedman, J. 2021. Autopolyploid establishment depends on life history strategy
724 and the mating outcomes of clonal architecture. *bioRxiv* 2021.10.21.465190.

725 Vanneste, K., Baele, G., Maere, S. & Van de Peer, Y. 2014a. Analysis of 41 plant genomes supports a
726 wave of successful genome duplications in association with the Cretaceous–Paleogene
727 boundary. *Genome research* **24**: 1334–1347. Cold Spring Harbor Lab.

728 Vanneste, K., Maere, S. & Van de Peer, Y. 2014b. Tangled up in two: a burst of genome duplications
729 at the end of the Cretaceous and the consequences for plant evolution. *Philosophical*
730 *Transactions of the Royal Society B: Biological Sciences* **369**: 20130353. The Royal Society.

731 Wang, W., He, Y., Cao, Z. & Deng, Z. 2018. Induction of tetraploids in impatiens (Impatiens
732 walleriana) and characterization of their changes in morphology and resistance to downy
733 mildew. *HortScience* **53**: 925–931. American Society for Horticultural Science.

734 Wood, T.E., Takebayashi, N., Barker, M.S., Mayrose, I., Greenspoon, P.B. & Rieseberg, L.H. 2009. The
735 frequency of polyploid speciation in vascular plants. *Proceedings of the national Academy of*
736 *sciences* **106**: 13875–13879. National Acad Sciences.

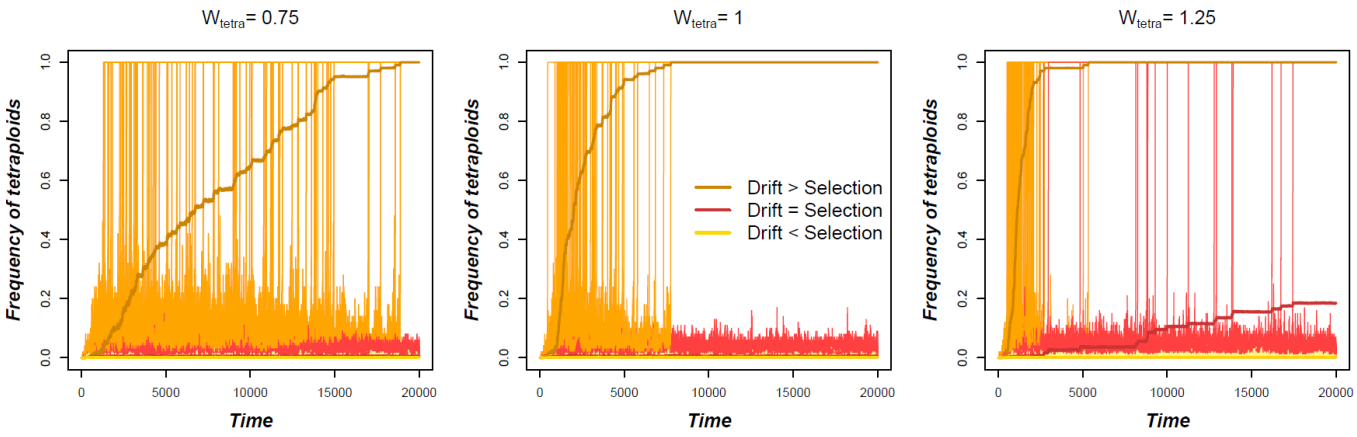
737 Yao, Y., Carretero-Paulet, L. & Van de Peer, Y. 2019. Using digital organisms to study the evolutionary
738 consequences of whole genome duplication and polyploidy. *PloS one* **14**: e0220257. Public
739 Library of Science San Francisco, CA USA.

740

741

742

743



744

745 **Figure 1.** Dynamics of the proportion of tetraploid individuals through time for different
746 strengths of genetic drift when tetraploid individuals are less fit (left panel), as fit (middle
747 panel) or fitter (right panel) than their diploid progenitors. The thin lines represent individual
748 simulations (n = 100), and the thick line is the mean proportion of polyploids among
749 simulations. Other parameters are L = 100 (and then $\bar{s_0}$ = 0.01), h = 0.5, σ = 0, U = 0.1 and N
750 can be equal to 50 (Drift > Selection), 100 (Drift = Selection) or 200 (Drift < Selection).

751

752

753

754

755

756

757

758

759

760

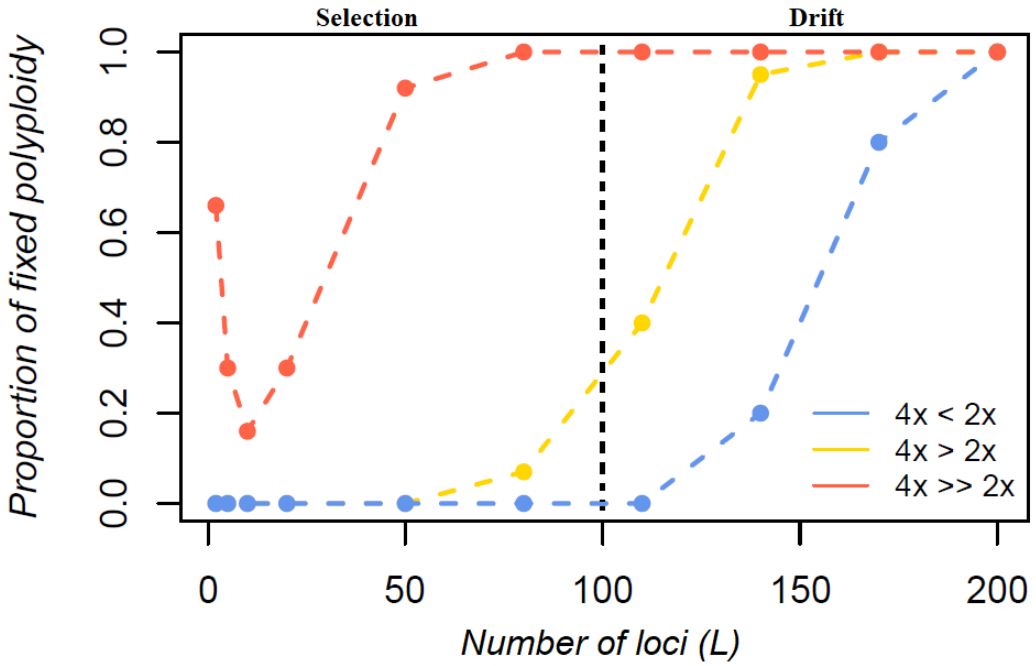
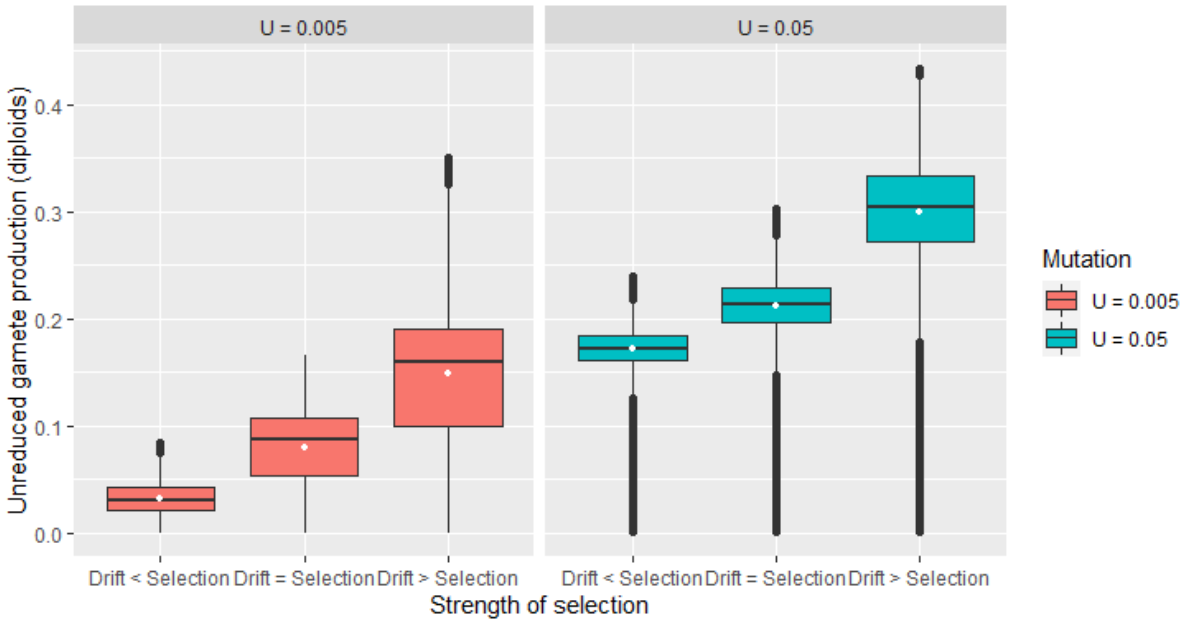


Figure 2. Probability of fixation of polyploidy as a function of the number of loci underlying the trait and the relative fitness of diploid and tetraploid individuals (W_{tetra} = 0.75 in blue, 1.25 in yellow, or 2 in red). N is equal to 100, and L varies from 2 to 200, genetic drift is stronger than selection when $L > 100$. Other parameters are $h = 0.5$, $\sigma = 0$, $U = 0.1$, $\bar{s_0}$ varies from 0.005 ($L = 200$) to 0.5 ($L = 2$). The probability is computed on $n = 100$ replicates.

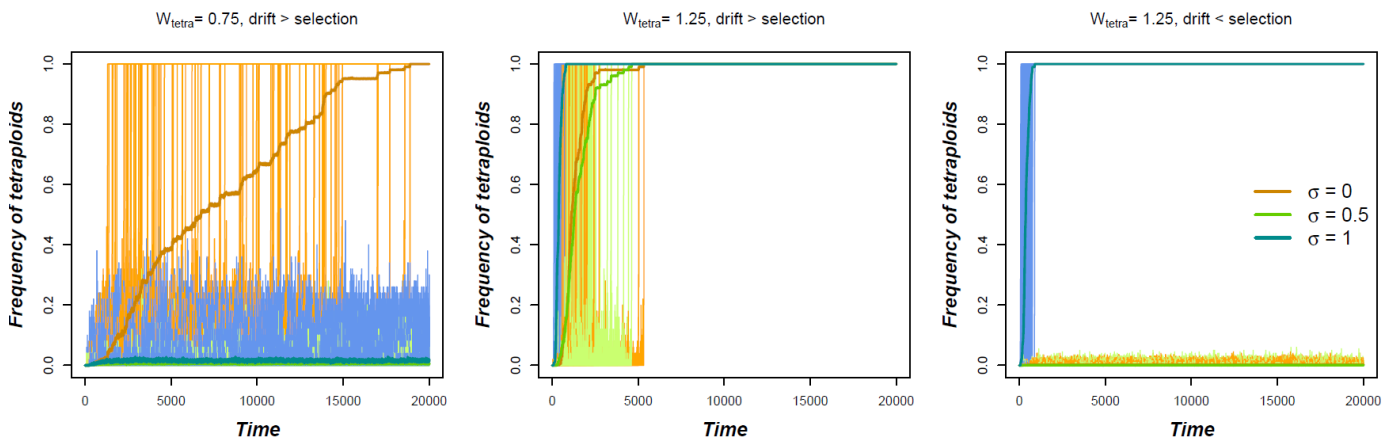


774

775 **Figure 3.** Production of unreduced gametes in diploid individuals, as a function of strength of
776 selection compared to genetic drift, for different mutation rates (left panel U = 0.005, right
777 panel U = 0.05), when tetraploid individuals are less fit than diploid ones. The white points
778 represent the average production of unreduced gametes and is estimated with respectively n =
779 5000, 10.000 and 20.000 estimates respectively when N = 50, 100 or 200 (100 replicates
780 times the population size). Other parameters are L = 100 (and then $\bar{s_0}$ = 0.01), h = 0.5, σ = 0
781 and N can be equal to 50 (Drift > Selection), 100 (Drift = Selection) or 200 (Drift <
782 Selection).

783

784



785

786 **Figure 4.** Dynamics of the proportion of tetraploid individuals in time, as a function of the
787 strength of the selfing rate in the population (σ = 0: obligate outcrossers in orange, σ = 0.5:
788 non-obligate outcrossers in light green, σ = 1: obligate selfers in blue). (Left panel) Polyploids
789 are less fit than diploids, and genetic drift is stronger than natural selection. (Middle panel)
790 Polyploids are fitter than diploids, and genetic drift is stronger than natural selection. (Right
791 panel) Polyploids are fitter than diploids, and genetic drift is weaker than natural selection.
792 The thin lines represent individual simulations (n = 100), and the thick line is the mean
793 proportion of polyploids among simulations. Other parameters are L = 100 (and then $\overline{s_0}$ =
794 0.01), h = 0.5, U = 0.1 and N can be equal to 50 (Drift > Selection) or 200 (Drift < Selection).

795