

SPECIAL ISSUE ARTICLE

Why are triploid quaking aspen (*Populus tremuloides*) common?

 Benjamin Wong Blonder 

Department of Environmental Science, Policy, and Management, University of California at Berkeley, Berkeley, CA 94720 USA,

Correspondence

 Benjamin Wong Blonder, Department of Environmental Science, Policy, and Management, University of California at Berkeley, Berkeley, CA 94720 USA.
 Email: benjamin.blonder@berkeley.edu

 This article is part of joint special issues of the *American Journal of Botany and Applications in Plant Sciences*: “Twice as Nice: New Techniques and Discoveries in Polyploid Biology.”

Abstract

Premise: Quaking aspen is a clonal tree species that has mixed ploidy, often with high relative abundance of both diploids and triploids but no haploids or tetraploids. Triploids typically have low fertility, leaving their occurrence apparently unlikely from an evolutionary perspective, unless they provide a “triploid bridge” to generating higher-fitness tetraploids—which are not observed in this species. This study focused on how triploidy can be maintained in quaking aspen.

Methods: A computational model was used to simulate gamete production, sexual reproduction, asexual reproduction, parent survival, and offspring survival in a population. All parameters were assumed to be cytotype-dependent and environment-independent. Sampling methods were used to identify parameter combinations consistent with observed cytotype frequencies.

Results: Many processes and parameter values were sufficient to yield a moderate frequency of triploids, and very few were necessary. The most plausible route involved higher triploid survival at the parent or offspring stage and limited unreduced gamete production by either diploid or triploid parents. Triploid fertility was helpful but not necessary.

Conclusions: The coexistence of diploids and triploids in quaking aspen is statistically likely and promoted by the existence of commonly observed, long-lived triploid clones. However, other mechanisms not captured by the model related to environmental variation could also occur. Further empirical data or more complex but difficult-to-parameterize models are needed to gain further insight.

KEYWORDS

mixed ploidy, polyploidy, quaking aspen, triploid bridge, triploid sterility, unreduced gametes

Many species exhibit intraspecific variation in cytotype (ploidy level) (Keeler, 1998; Fox et al., 2020). This variation can sometimes lead to speciation (Otto, 2007), but can also be temporally persistent. Intraspecific ploidy levels are typically observed to be even-numbered ($2n = 2x, 4x, 6x$, etc., where n is the typical gametic chromosome number and x is the base number of chromosomes). Odd-numbered ploidies (e.g., $2n = 3x$, triploids) are typically rare. A common assumption is that such odd-numbered ploidies have varying but typically low fertility due to meiotic failures in which the low probability of successful chromosome segregation into reduced gametes ($n = x$) results in aneuploidy and prezygotic barriers to reproduction (Jackson, 1976; Zielinski and Mittelsten Scheid, 2012). Additionally, postzygotic barriers can also occur via a “triploid block” (Köhler et al., 2010), whereby unbalanced

gene product ratios between parental genomes lead to low performance. Odd-numbered ploidies are thought to be unlikely to evolve in dioecious species where reproductive problems occur due to disruptions in gene product ratios between (potentially incipient; Hou et al., 2015) sex chromosomes and autosomal chromosomes (Marks, 1966).

Nevertheless, odd-numbered ploidies are sometimes observed within species at high frequencies and form via intraspecific mechanisms. A non-comprehensive list of intraspecific examples includes *Andropogon* (Poaceae) (Keeler, 1990), *Arabidopsis* (Brassicaceae) (Henry et al., 2005), *Campuloclinium* (Asteraceae) (Farco and Dematteis, 2014), *Chamaerion* (Onagraceae) (Burton and Husband, 2001), *Boechera* (Brassicaceae) (Schranz et al., 2005), *Dryopteris* (Dryopteridaceae) (Lin et al., 1992), *Ginkgo* (Ginkgoaceae) (Šmarda et al., 2018), *Solidago* (Asteraceae) (Schlaepfer et al., 2008), and *Triticum*

(Poaceae) (Namikawa and Kawakami, 1934). Odd-numbered ploidy hybrids that formed via interspecific mechanisms are not discussed further here (e.g., in *Betula* [Betulaceae]; Anamthawat-Jónsson and Thór Thórsson, 2003).

The persistence of odd-numbered ploidies in populations at high frequencies therefore requires addressing the question of stable coexistence. How can they coexist with diploids and individuals with other even-numbered ploidies, if they have low fertility and if minority cytotype exclusion (Levin, 1975) prevents their establishment from low initial frequencies?

This study addresses this general cytotype coexistence question in quaking aspen, *Populus tremuloides* Michx. (Salicaceae), a wide-ranging North American tree species (Mitton and Grant, 1996). This dioecious species occurs as both diploids and triploids of both sexes (Van Buijtenen, 1957). Haploids have not been observed (Mock et al., 2012), and tetraploids are probably rare (Einspahr and Winton, 1977). Unreduced gametes (pollen) are frequently observed (Benson and Einspahr, 1967; Winton and Einspahr, 1970). Both diploid and triploid cytotypes co-occur at local and regional scales (Mock et al., 2008; Blonder et al., 2021) (Figure 1). Triploids are more common in the southwestern range (Mock et al., 2012; Goessen et al., 2022). Triploids can comprise approximately 50% of the individuals sampled in some regions such as southwestern Colorado (Blonder et al., 2021). Notably, the species also grows clonally (Mitton and Grant, 1996). Triploids are also more likely to form large spatially extensive clones (Mock et al., 2012). For example, the “Pando” clone is a triploid individual that comprises ~40,000 stems over 43 ha of land (DeWoody et al., 2008).

Aspen is a useful test case for understanding odd-numbered ploidy coexistence because its triploids pose an apparent paradox (described below). It is also an interesting

test case for the more general question of the evolution and coexistence of odd-numbered ploidies in mixed-ploidy populations because assessments of this question in other species or in general models have typically focused on scenarios where triploids are rare (Husband, 2004).

On the one hand, triploids comprise large fractions of the population in many environments, when counted either by genotype or by land area (Blonder et al., 2021; Goessen et al., 2022). They also have high vegetative growth rates (Benson and Einspahr, 1967; Every and Wiens, 1971; DeRose et al., 2015; Blonder et al., 2021). On the other hand, triploids are said to have low fertility (DeRose et al., 2015; Greer et al., 2017) or the question of low fertility is acknowledged but unresolved (Mock et al., 2012). Limited fertility of natural triploids has been observed (e.g., Fig. S29 of Goessen et al., 2022; K. Mock, Utah State University, personal communication). However, triploid mothers that do flower often do not develop seeds or develop few seeds that reach maturity. Of these seeds, germination rates tend to be low and resulting seedlings often have growth defects. However, germinated seeds of triploid mothers often are triploids, while no tetraploid mothers have been observed to produce seeds (K. Mock, Utah State University, personal communication). Limited triploid fertility (and high subsequent offspring mortality) also has been reported in the related *P. tremula* (e.g., Fig. 15 of Johnsson, 1940).

The apparent paradox is then: How can quaking aspen populations have high frequencies of diploids and triploids and low or zero frequencies of haploids and tetraploids if triploids have low fertility? There are at least 11 hypotheses to consider (Table 1) (see also the broader discussion of Ramsey and Schemske, 1998). These hypotheses generalize beyond quaking aspen. They are also not mutually exclusive, meaning multiple mechanisms may contribute

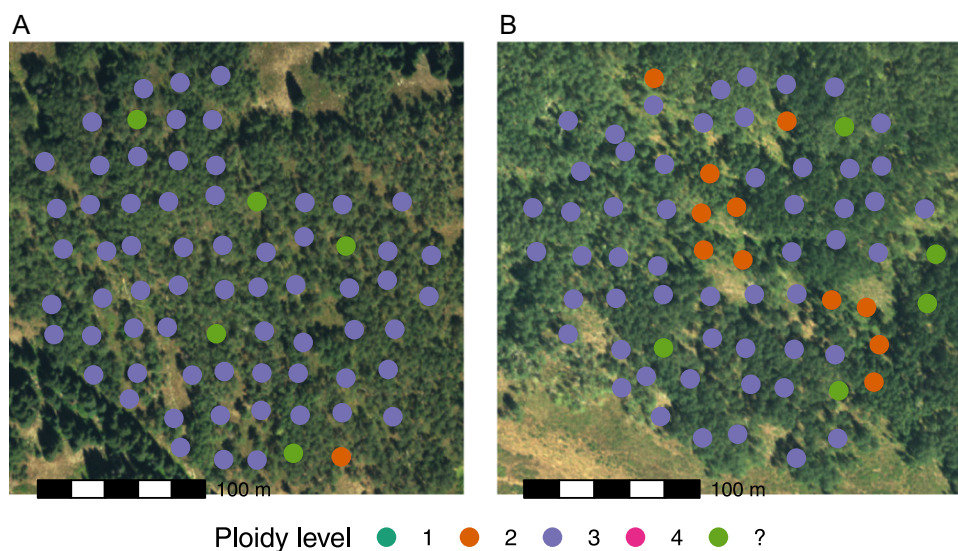


FIGURE 1 Example of co-occurrence of diploid and triploid quaking aspen. Points indicate individual stems and are plotted over a satellite base map, with cytotype indicated by point color. Data are re-plotted from gridded genetic sampling within areas in southwestern Colorado originally reported by Blonder et al. (2021). (A) Coal Creek, (B) Middle East River.

TABLE 1 Symbols used within the model.

Type	Symbol	Interpretation	Type
Model output	d	Distance between model output and target output	Scalar
Model output	\bar{F}	Mean frequency of each cytotype over a time interval	Vector of length 4
Model output	$\vec{F}(t)$	Relative frequency of each cytotype at time t	Vector of length 4
Parameter	M	Population size	Scalar
Parameter	A	Probability a parent of a certain cytotype produces a gamete of a certain cytotype	Matrix of size 4×5
Parameter	\vec{B}	Probability a parent of a certain cytotype produces an offspring of the same cytotype via apomixis	Vector of length 4
Parameter	\vec{C}	Probability an offspring of a certain cytotype survives to become a parent	Vector of length 4
Parameter	\vec{D}	Probability a parent of a certain cytotype survives to the next generation	Vector of length 4
Parameter	\vec{F}_{target}	Desired relative frequencies of cytotypes in model output	Vector of length 4
Parameter	g	Number of gametes produced per parent	Scalar
Parameter	τ	Interval of time used to average final model output	Scalar
Parameter	t_{max}	Number of timesteps to iterate the model	Scalar
Variable	t	Current timestep	Scalar
Variable	$\vec{P}(t)$	Cytotype of each parent individual at time t	Vector of length M
Variable	$\vec{G}(t)$	Cytotype of each gamete at time t	Vector of length in range $[0, gM]$
Variable	$\vec{O}_{\text{sexual}}(t)$	Cytotype of each sexually produced offspring at time t	Vector of length in range $[0, gM]$
Variable	$\vec{O}_{\text{asexual}}(t)$	Cytotype of each asexually produced offspring at time t	Vector of length in range $[0, gM]$
Variable	$\vec{O}(t)$	Offspring at time t	Vector of length in range $[0, 2gM]$
Variable	$\vec{O}_s(t)$	Surviving offspring at time t	Vector of length in range $[0, M]$
Variable	$\vec{P}_s(t)$	Surviving parents at time t	Vector of length in range $[0, M]$

simultaneously to observed patterns. (1) Odd-numbered ploidies could have similar fitness to even-numbered ploidies due to variation in any fitness component, e.g., growth, survival or fecundity (gamete and seed production). Trade-offs among fitness components could also exist, such that higher survival could compensate for lower fecundity. In aspen, this hypothesis is plausible given the high growth rates and vegetative dominance of triploids. (2) Gamete production rate could vary; i.e., each parent cytotype produces gametes of different cytotypes at different rates (e.g., if viable reduced gametes are produced at a high rate or if viable unreduced gametes are produced at a high rate). Odd-numbered ploidy offspring could be generated if even-numbered ploidies have high fitness and produce reduced gametes that yield odd-numbered ploidy offspring (i.e., the first but not second stage of the “triploid bridge”). In aspen, this hypothesis is plausible given the frequent observation of unreduced gametes. (3) A “triploid bridge” could exist (Husband, 2004) in which triploids are initially produced by tetraploid and diploid parents producing reduced gametes,

yielding triploid offspring with low fitness; these offspring then could produce unreduced triploid gametes and combine with reduced gametes of diploid parents to produce tetraploid offspring, yielding positive feedback (and similarly for other odd-numbered ploidies). If the tetraploid offspring have sufficiently high fitness, then the triploid individuals may be maintained in the population despite low fitness. In aspen, this hypothesis is probably not plausible given the absence of tetraploids. (4) Apomixis could occur, in which asexual reproduction without fertilization occurs (Asker and Jerling, 1992), here for odd-numbered ploidies (e.g., triploid parent asexually producing triploid offspring). In aspen, this hypothesis is probably implausible given the absence of evidence for apomixis despite investigation (Nagaraj, 1952). (5) Vegetative reproduction (clonal growth) could preferentially occur for odd-numbered ploidies, effectively reflecting either apomixis or higher parent survival. Clonal growth is common in aspen and seems to be more common in triploids (Mock et al., 2008, 2012). (6) Drift could occur, in

which odd-numbered and even-numbered ploidy levels have equivalent fitness and shift in abundance stochastically. For aspen, the amount of data available is insufficient to assess this possibility. (7) Odd-numbered ploidy levels could occur in another population for whatever reason and disperse into the local population at high rates. For example, polyploidy is thought to confer performance advantages compared to diploids in stressful environments (Van de Peer et al., 2021); thus, odd-numbered ploidy levels may be found in non-stressful environments where they have lower fitness than diploid individuals but persist as “sink” populations if there is high propagule pressure (Kawecki, 2004). In aspen, this hypothesis is plausible because triploids are thought to occur primarily in drier and warmer environments, reflecting a potential spatial niche partitioning necessary for this phenomenon (Mock et al., 2012; Blonder et al., 2021). (8) A competition–colonization trade-off could occur, if there were differences in the dispersal and establishment rates of odd- and even-numbered ploidy levels and multiple habitat patches were available. This hypothesis is also plausible in aspen given the higher growth rates of triploids (Benson and Einspahr, 1967). (9) If temporal environmental variation occurs, then a storage effect or relative nonlinearity of competition (Chesson, 2000) could lead respectively to the favoring of odd-numbered ploidy levels in certain rare environments or in temporally variable environments. This hypothesis is plausible in aspen given the potentially long lifespan of clones relative to interglacials (DeWoody et al., 2008) and the capacity of the species for strong reproductive responses to disturbance events (DeByle and Winokur, 1985). (10) Lags could occur, such that historical environments selected for the high prevalence of odd-numbered ploidy, but current environments are now selecting against them. The observed presence of odd-numbered ploidy levels would then be transiently dynamic. In aspen, this hypothesis is plausible because triploids are currently declining in drier and warmer environments (Blonder et al., 2021) and many aspen forests are undergoing successional dynamics due to prior settler-colonial land disturbance or shifts in fire regime (Rogers et al., 2014). (11) A form of intraspecific parasitism could occur, such that odd-numbered ploidy levels, especially those representing large clones, have negative effects on population fitness. Parasitism could occur either directly through resource preemption as large odd-numbered ploidy clones dominate a spatial extent, precluding reproductive even-numbered ploidy individuals from establishing, or indirectly through reducing standing genetic variation (Fisher, 1930) in temporally variable environments. However, such intraspecific parasitism could also be selected for in the long term if resource preemption strengthens interspecific competition, i.e., preventing other species from accessing preferred habitat. This hypothesis is plausible in aspen given its frequent occurrence as monodominant stands (DeByle and Winokur, 1985).

Given that data are highly limited, this study used a modeling approach to make inferences about these hypotheses. The study focused on assessing Hypotheses 1–6 because they are spatially and temporally implicit and

can be compactly modeled. Assessing Hypotheses 7–11 would require a model that was more comprehensive, and more difficult to parameterize and falsify; they also would require more detailed exploration of *stable* coexistence and therefore the relative strengths of intra- vs inter-cytotype competition (Chesson, 2000) than is of interest here. This study thus represents a first effort toward the broader question.

First, a general mathematical model was developed for the relative frequency of different cytotypes in a population that experiences iterative sexual and asexual reproduction events, inspired by similar models of Husband (2004) and Yamauchi et al. (2004). The resulting model has many parameters, for which most of the values are unknown. Model predictions were then generated across the entire parameter space, and predictions were classified as rejected or accepted based on the observed distribution of cytotypes. Because parameters map onto hypotheses, the accepted parameter space then allows inference about which hypotheses may be consistent with observations.

MATERIALS AND METHODS

Model overview and assumptions

This model is spatially implicit and includes no environmental change or evolutionary shifts in parameters over time. It simulates parent fertility (gamete production), sexual reproduction, apomixis, offspring survival, and parent survival within a population comprising different cytotypes, with all parameters cytotype-dependent. For aspen, only haploid, diploid, triploid, and tetraploid cytotypes were considered. Clonal growth can occur via the apomixis or parent survival parameters; distinguishing among them is not possible in the model. Sexual and/or asexual reproduction occurs in non-overlapping generations and is assumed to be random with respect to cytotype. The population is assumed to have a fixed and finite size, with no immigration or emigration. The model is iterated over time until a quasi-equilibrium is reached.

Individuals in the model are best interpreted as ramets (not genets) in this species. Because individuals are distinguished only by their cytotype (not genotype) and are constrained to a fixed population, the model can be interpreted as representing competition among stems in a habitat of fixed size. Because genotype does not come into the model, clone size is not explicitly represented, except in that high abundance of a certain cytotype is a necessary prerequisite for the existence of a large clone.

Model structure and dynamics

All symbols used in this section are summarized in Table 1.

The parent individuals in the population at time t are represented as a vector with entries for each cytotype,

$\vec{P}(t) = \{P_i\}$, $1 \leq i \leq M$ where M is the total number of individuals in the population, and P_i can take values from 1 to 4 corresponding to each cytotype.

At each integer time t , the following processes occur:

- (1) Gamete production. Each parent individual produces viable gametes. The probability that a parent of cytotype p_i produces a viable gamete of cytotype j is given by the matrix A with entries A_{ij} where $1 \leq i \leq 4$ and $0 \leq j \leq 4$; the first column corresponds to the probability of producing an empty gamete with zero ploidy level. A parent is assumed to produce up to g gametes. The gamete production process is carried out through random sampling for each parent with weights taken from A and yields a vector $\vec{G}(t)$ with possible values from 0 to 4 and length in the range $[0, gM]$.
- (2) Sexual reproduction. Two gametes k and l are selected at random, without replacement, for sexual reproduction (selfing is allowed, but becomes vanishingly unlikely as M becomes large). They produce an offspring vector $\vec{O}_{\text{sexual}}(t) = \{O_{\text{sexual},i}(t)\}$ of length M , each of whose entries has possible values from 1 to 4 and equal to $\vec{O}_{\text{sexual}}(t) = G_k(t) + G_l(t)$ (i.e., the offspring cytotype is the sum of each parent's gamete cytotype). This process is repeated for all gamete pairs. Any offspring with $O_i(t) > 4$ (i.e., pentaploid or higher) or $O_i(t) = 0$ (i.e., no genome) is removed from the population. While this removal represents a potential bias toward lower ploidy levels, the lack of observed higher ploidy levels in this species makes this a reasonable assumption. The output vector has a length in the range $[0, gM]$.
- (3) Asexual reproduction (apomixis). Each parent independently is allowed to asexually produce offspring of the same cytotype (e.g., vegetative/clonal growth). The apomixis process is carried out through random sampling of g offspring for each parent with weights taken from a vector \vec{B} with entries B_i where $1 \leq i \leq 4$. The outcome of this process is a vector $\vec{O}_{\text{sexual}}(t) = \{O_{\text{asexual},i}(t)\}$ with possible values from 1 to 4 and length in the range $[0, gM]$.
- (4) Offspring survival. A final vector of both types of offspring is constructed, $\vec{O}(t) = \{O_{\text{asexual},i}(t), O_{\text{sexual},i}(t)\}$. Offspring are then assumed to survive to parenthood dependent on cytotype. The probability of this occurring for cytotype i is given by a vector \vec{C} with entries C_i where $1 \leq i \leq 4$. This parameter is interpretable as an integration of life cycle components from embryo survival, seed production, seedling establishment, and adult survival and does not distinguish among sexual and asexual offspring. Each offspring independently is allowed to survive, or not. Dead individuals are removed.

- (5) Parent survival. Parents are allowed to survive into the next generation (reflecting differential lifespans and potential clonal growth). The probability of parent survival for cytotype i is given by a vector \vec{D} with entries D_i where $1 \leq i \leq 4$. Each parent independently is allowed to survive, or not. Dead individuals are removed, yielding $\vec{P}_s(t)$. The difference between the number of surviving parents and M is used to select the number of surviving offspring that contribute to the final population. The surviving offspring are resampled with replacement to the appropriate total, yielding $\vec{O}_s(t)$. The parent and offspring survivors are then combined to yield the next generation, $\vec{P}(t+1) = \{\vec{O}_s(t), \vec{P}_s(t)\}$.

This process is iterated for t_{max} timesteps, at each of which $t \rightarrow t+1$. At each timestep, the relative frequency of each cytotype is calculated, as $\vec{F}(t) = \{F_j(t)\} = \{(\sum_i P_i(t) = j)/M\}$, for $1 \leq j \leq 4$. The final statistic used for statistical inference is \bar{F} , the mean value of $\vec{F}(t)$ over an interval of length τ , i.e., from $(t_{\text{max}} - \tau) \leq t \leq t_{\text{max}}$. Similar statistics are calculated for other model variables that change over time.

Model scenarios

The study considered scenarios in which five different subscenarios were toggled on/off computationally, yielding $2^5 = 32$ total scenarios. Each scenario was explored using random sampling of parameters. This approach was taken, rather than nonlinear optimization, to understand the full landscape of model predictions and because full enumeration of the parameter space was infeasible. Additionally, to reduce the problem space dimensionality, it was assumed that $A_{i,j} = 0$, $\forall j > (i+1)$, i.e., gametes cannot increase cytotype beyond their parent's value, but can have the same, lower, or zero cytotype, meaning the dimensionality of the problem over the $(A, \vec{B}, \vec{C}, \vec{D}, M)$ parameter space is $(2+3+4+5)+4+4+4+1=23$.

The scenarios were intended to provide counterfactual assessments of the importance of certain processes. The first subscenario, 'include_apomixis' if 'off' set $\vec{B} = 0$. The second subscenario, 'include_offspring_survival_variation', if 'off' set $\vec{C} = 0$. The third subscenario, 'include_haploid_tetraploid', if 'off' set several parameters to zero: no gamete production from haploids ($A_{1,j} = 0$) or tetraploids ($A_{4,j} = 0$), no apomixis of haploids ($B_1 = 0$) or tetraploids ($B_4 = 0$), no offspring survival of haploids ($C_1 = 0$) or tetraploids ($C_4 = 0$), and no parental survival of haploids ($D_1 = 0$) or tetraploids ($D_4 = 0$). The fourth subscenario, 'include_triploid_fertility', if 'off' set gamete production for triploids to zero ($A_{3,1}, A_{3,2}, A_{3,3}, A_{3,4} = 0$). The fifth subscenario, include_parent_survival, if 'off' set $\vec{D} = 0$.

Otherwise, all parameters were allowed to vary freely. For A , values of each row must have a sum between 0 and 1 (not necessarily exactly 1). For each row's values, excluding those values assumed to be zero (by the assumption in the

above paragraph), values were sampled uniformly randomly between 0 and 1, then normalized by their row sum, and then multiplied by another uniformly random sampled number between 0 and 1. For each of \vec{B} , \vec{C} , and \vec{D} , values were independently drawn for each entry from the uniform distribution between 0 and 1. For M , values were uniformly randomly sampled from 10 to 1000.

Model runs

A total of 20,000 samples were obtained for each scenario. The population was assumed to be all-diploid initially, i.e., $p_i(t=0) = 2, \forall i$, and a maximum gamete count of $g=10$ (this parameter affects computational runtime only). Each iteration was run for $t_{\max} = 100$ timesteps, with final results averaged over $\tau = 10$ timesteps. Based on time-series plots, this duration was sufficient to reach a quasi-steady state in nearly all cases (Figure 2).

Model run acceptance

The target of the model was set to be a composition of $\vec{F}_{\text{target}} = \{0,0.75,0.25,0\}$, i.e., 75% diploid, 25% triploid.

This value approximately reflects observed frequencies in western North America (Mock et al., 2012; Blonder et al., 2021; Goessen et al., 2022) and the likely infrequent but potentially nonzero occurrence of haploids and tetraploids (Van Buijtenen, 1957; Einspahr et al., 1963). The qualitative conclusions of the study do not have a strong dependence on the precise values of \vec{F}_{target} used (data not shown), and in any case, there is likely substantial variation in triploidy frequency across locations. For each simulation run, the distance from the target (i.e., quadratic error) was calculated as $d = \sqrt{\sum_{i=1}^4 (\bar{F}_i - \bar{F}_{\text{target},i})^2}$. A simulation run was considered “in-target” according to the criterion $d < 0.1$.

Statistical analyses

The overall goal was to determine which parameters in the $(A, \vec{B}, \vec{C}, \vec{D}, M)$ space predicted a simulation run being in-target. A machine learning algorithm was used to identify the combinations of predictors associated with low values of distance, i.e. being in-target, for each scenario. A regression random forest was used, with d as the response variable and $(A, \vec{B}, \vec{C}, \vec{D}, M)$ as the predictors. This approach allows for nonlinear and interactive effects between predictors and allows us to approximate the response surface of the

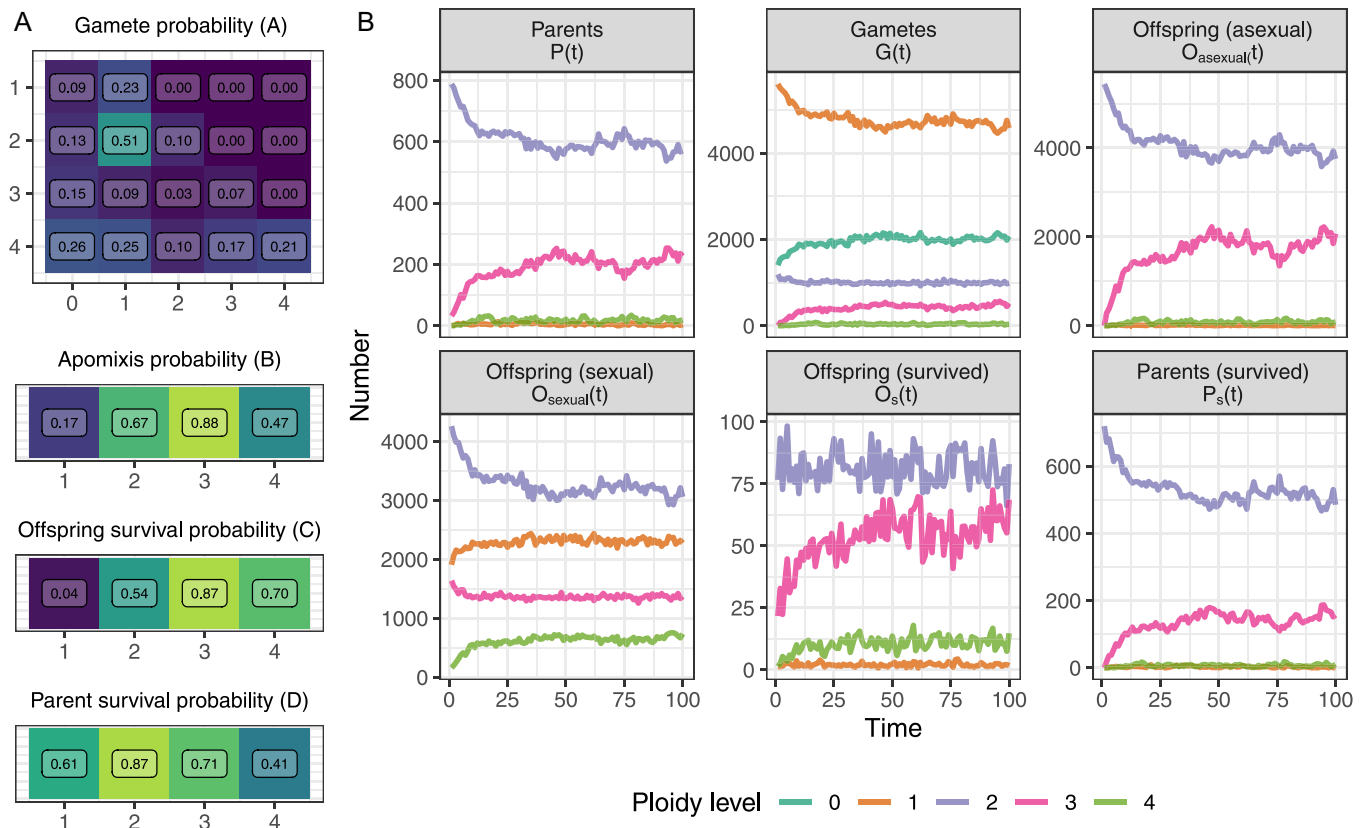


FIGURE 2 Example time series predicted by the model in the scenario combining all subscenarios for a randomly selected outcome where diploids and triploids co-occur at similar frequencies. (A) Panels show colored parameter values (M is not shown). (B) Model predictions. Lines are colored by cytotype.

simulation model without further computational expense (enumerating its parameter space or further random sampling of it). The number of variables to split at each node was set to 5, the number of trees to 1000, the minimum node size to 5, and the maximum depth to infinity. Because of the low number of cases that were in-target, a balancing procedure was carried out. Case weights were set to the inverse relative prevalence of in-target cases for in-target cases and 1 otherwise.

Model fit was assessed with R^2 . Variable importance was assessed via a Gini index impurity score (Breiman, 2001). Based on these importance scores, variable combinations were plotted to assess their effect on d and to determine whether they represented biologically plausible cases. The independent effects of additional predictors were also assessed with partial dependence plots.

Simulations were implemented using the R package `tidyverse` (Wickham et al., 2019) in R version 4.2.0 (R Core Team, 2022). Random forests were fit and analyzed using R packages `ranger` (Wright and Ziegler, 2017) and `pdp` (Greenwell, 2017).

RESULTS

Example time series of simulation predictions and their corresponding parameters are shown in Figure 2. This example is representative of the typical equilibration rate and level of stochasticity observed in the model. Time series outputs for the full set of simulations are also archived. A full exploration of all scenarios and parameters is not possible for space reasons, so the text below focuses on the most interesting cases.

Many but not all scenarios and parameters yield plausible outcomes

There was high variability in the cytotype frequencies predicted by the model, both within and across scenarios (Figure 3). Across scenarios, outcomes tended to be dominated either by primarily diploid/triploids, or by haploids/tetraploids, with fewer outputs corresponding to even mixtures of all cytotypes. The overall rate of in-target simulations across the parameter space ranged from 0% to 19% across scenarios (Appendix S1: Figure S1), or after normalization to account for differing dimensionalities, 0 to 84%. The underlying distance-to-target distributions are visualized in Appendix S1 (Figure S2). The lowest normalized rates (i.e., least evolvable scenarios) were the “none”, “triploid fertility”, “triploid fertility, haploid tetraploid”, and “triploid fertility, apomixis, haploid tetraploid” scenarios, all with rates below 20%.

The more common higher rates indicate that most scenarios, and also numerous parameter combinations within each, are sufficient to yield in-target outcomes. Notably, if other processes independently operate, triploid fertility is not

required, nor is the occurrence of haploid and tetraploid cytotypes, nor is variation in parent or offspring survival, or apomixis. However, variation in gamete production variation alone was not sufficient to yield in-target outcomes. Nevertheless, some of these in-target outcomes may rely on actual parameter values that are unrealistic, so more detailed assessment was also carried out.

Some parameters are more influential than others

Density plots of each parameter's marginal distribution indicated that some parameters likely had negligible effects on in-target outcomes, while others had large effects. Results were largely consistent across scenarios; see Figure 4 for the full scenario and Appendix S1 (Figure S3) for all other scenarios. Notably, M had a negligible effect, indicating that drift is not a key process influencing in-target outcomes. The gamete production (A) and apomixis parameters (\vec{B}) had generally smaller effects, except in some scenarios, while the offspring survival (\vec{C}) and parent survival (\vec{D}) parameters showed larger and consistent shifts in all scenarios. Note that a shift in marginal distribution is a sufficient but not necessary condition for detecting a variable effect.

Random forest models had reasonable success summarizing simulations, with R^2 values ranging from 10% to 87% across scenarios (mean = 64%, Appendix S1: Figure S4). Variable importances from these models yielded similar inferences. Figure 5 shows the full scenario and Appendix S1 (Figure S5) shows all other scenarios. Population size (M) had negligible impact on results in all scenarios. In the scenarios that included offspring survival (\vec{C}) or parent survival (\vec{D}), these parameters consistently had high importance. In scenarios including apomixis (\vec{B}) these parameters had moderate importances. Gamete transition probabilities (A) were typically least important. However, in a small number of scenarios (“none”, “offspring survival variation, parent survival”, “haploid tetraploid, parent survival”, “apomixis, haploid tetraploid, offspring survival variation”), these parameters emerged as being most important, particularly unreduced diploid gamete production rates (A_{22}).

Triploid offspring and parent survival comprise the main trade-off axis

The variable importances highlighted key parameters to investigate further. In the full scenario, plotting the ratio between triploid and diploid offspring survival vs. the ratio between triploid and diploid parent survival showed a strong negative relationship for the in-target outcomes (Figure 6A). That is, an in-target outcome can be achieved either by having relatively higher triploid offspring survival or relatively higher triploid parent survival but not both. These outcomes were consistent with a range of triploid

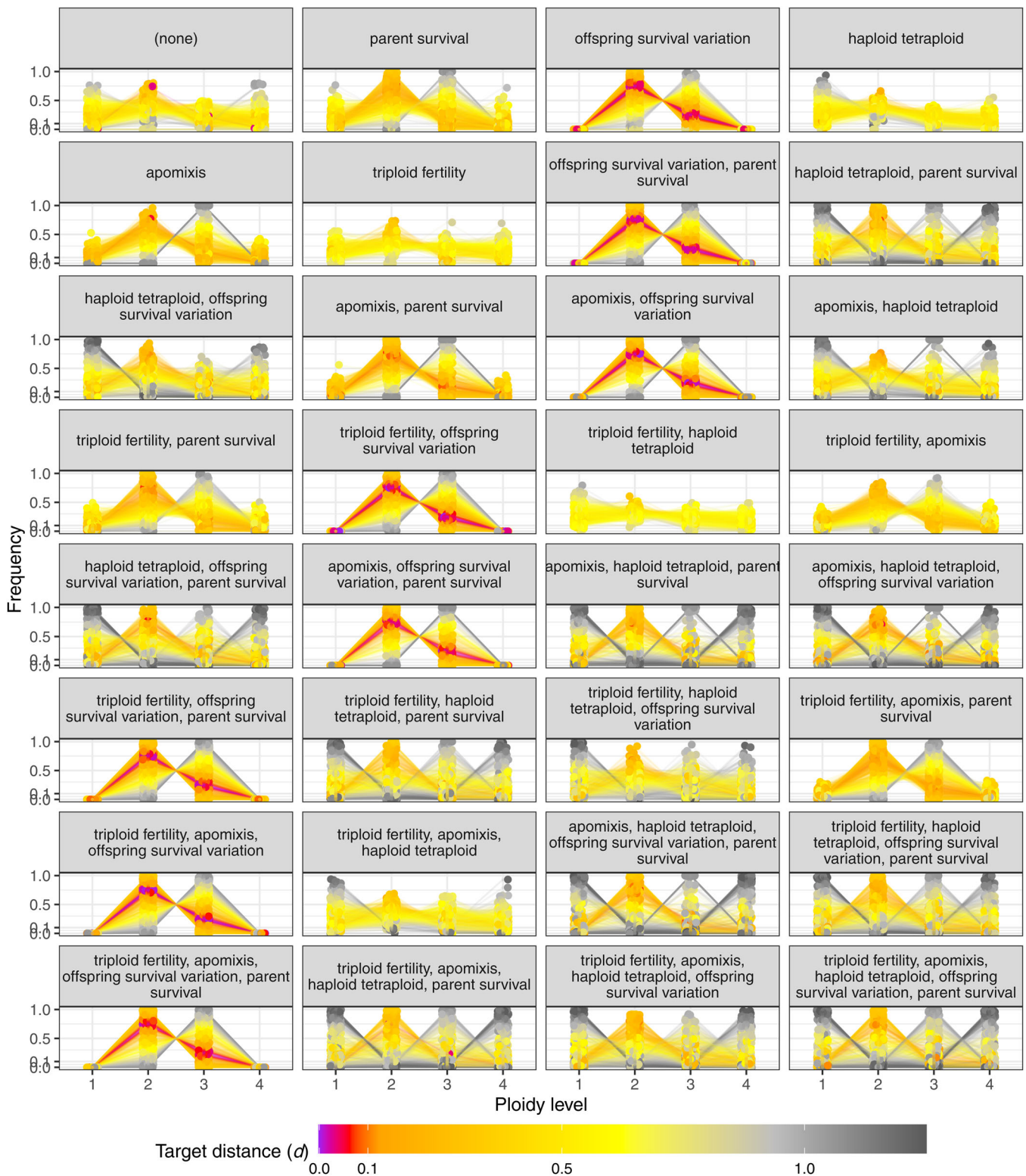


FIGURE 3 Distribution of cyotype frequencies (\bar{F}) predicted by the model across all sampled parameters under each scenario. Each simulation instance is connected by a line and colored by distance from target (d); in-target samples yielding $d < 0.1$ are shown in purple-red.

fecundities (Figure 6B), indicating that there was no survival–fecundity trade-off. However, some non-zero values are necessary to generate triploids through sexual reproduction, as the initial population is wholly diploid.

The strength of the offspring and parent survival trade-off was increased when scenarios were considered that eliminated the possibility of apomixis and haploid and tetraploid cyotypes (Figure 6C); additionally, a wider range

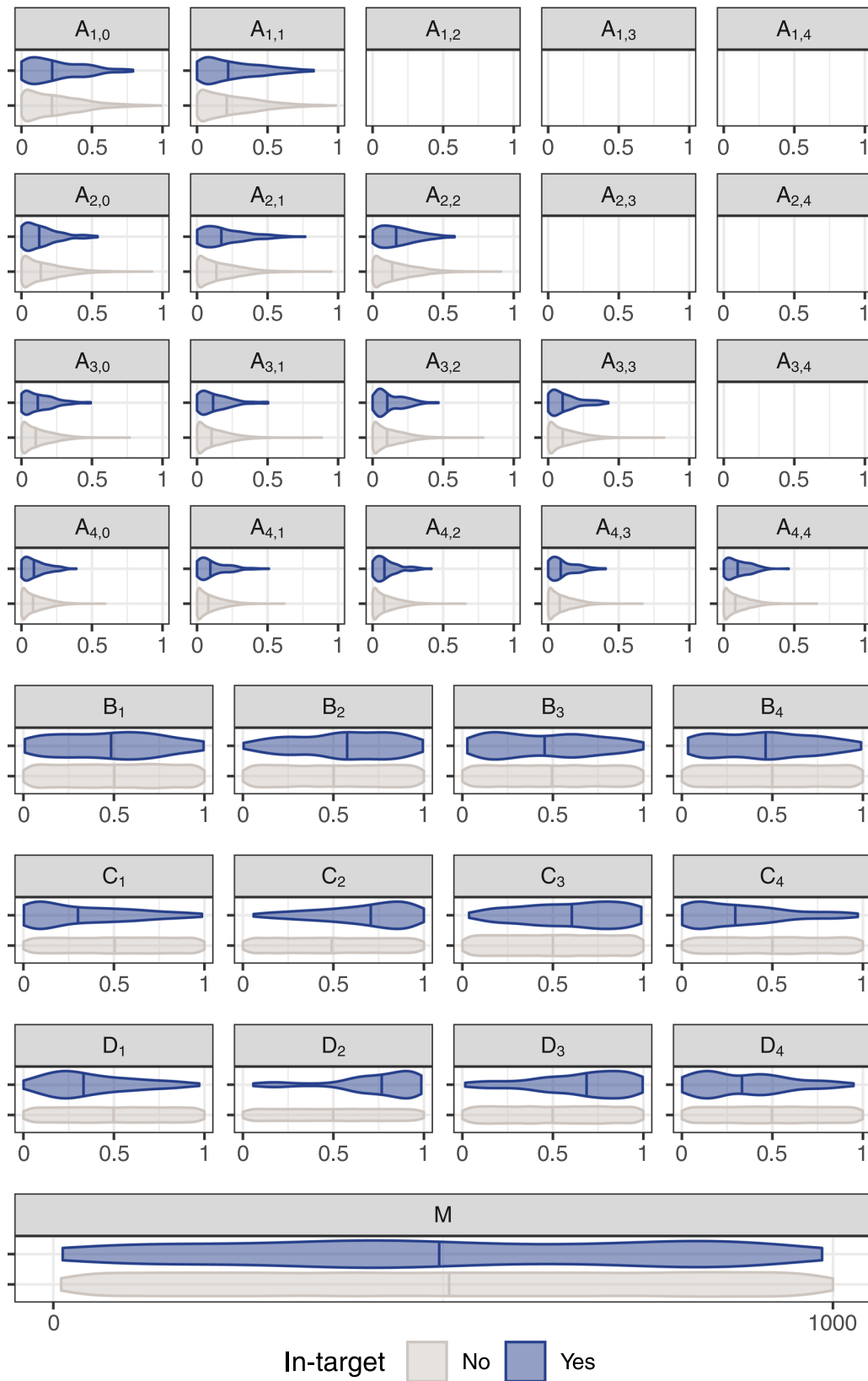


FIGURE 4 Marginal distributions of each parameter under the triploid fertility, apomixis, haploid tetraploid, offspring survival variation, parent survival scenario. Violin plots show medians as vertical lines and are colored by whether the cases are in-target ($d < 0.1$) or not. Results for the other scenarios are shown in Appendix S1 (Figure S3).

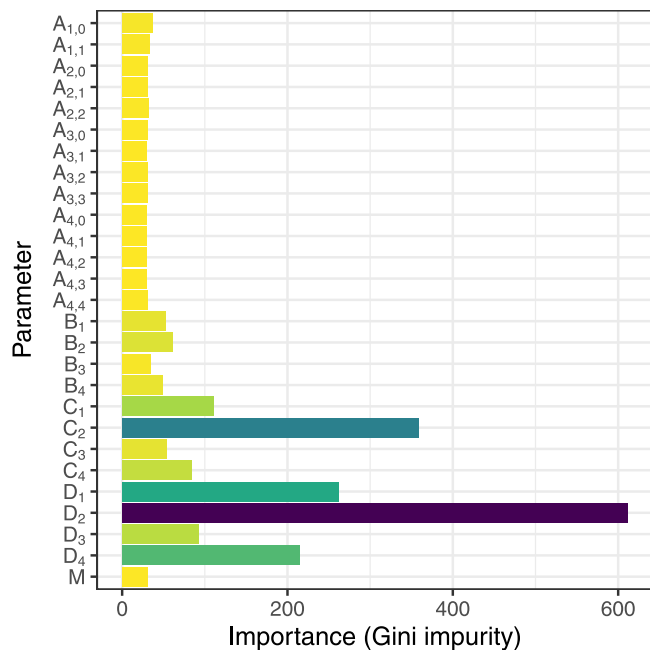


FIGURE 5 Variable importance scores (in random forest models) for each parameter in the triploid fertility, apomixis, haploid tetraploid, offspring survival variation, parent survival scenario. Larger bars or darker colors reflect higher importance. Results for the other scenarios are shown in Appendix S1 (Figure S4).

of survival combinations yielded in-target outcomes. The absence of a survival–fecundity trade-off for triploids also became more apparent (Figure 6D).

Unreduced gamete production or triploid apomixis could be important, but not triploid fecundity or haploid or tetraploid cytotypes

In the base scenario where only gamete production occurs, in-target outcomes occurred primarily for a narrow band of ratios between the probability of reduced and unreduced gamete production in diploid parents (Figure 7A). This result indicates the importance of unreduced gamete production, at an appropriate ratio, for in-target outcomes. Relationships for other gamete production probabilities were not as clear (not shown, data on Zenodo).

Notably, there was little effect of triploid fecundity on in-target outcomes (Figure 7B), indicating that gamete production by triploids is not necessary for in-target outcomes; rather, the unreduced gamete production by diploids is sufficient to maintain triploids.

Adding haploid and tetraploid cytotypes to the base scenario did not reveal an important role for reduced gamete production by tetraploids (Figure 7C), nor trade-offs between this rate and unreduced gamete production rate by diploids. This result indicates that haploid and tetraploid cytotypes do not substantially influence gamete production by diploids for in-target outcomes.

Adding apomixis to the base scenario did reveal additional trade-offs between relative apomixis probability of triploids relative to diploids and between gamete production of reduced gametes relative to unreduced gametes for diploids (Figure 7D). This result indicates in-target outcomes would require either higher triploid apomixis and low rates of unreduced gamete production in diploids or low triploid apomixis and high rates of unreduced gamete production in diploids.

Estimating gamete production rates

For better understanding the effects of each gamete production probability, partial dependence plots from the fitted random forest models were used to isolate the direct effect of each of these predictors (Appendix S1: Figure S6a) and identify the parameter values that yielded the lowest values of target distance. Notably, the best value for the probability a diploid parent produces an unreduced diploid gamete, A_{22} , was estimated at 0%, but with a range from 0% to 15% yielding values within a 50% relative range. The best value for the probability a triploid parent produces a reduced haploid gamete, A_{31} , was estimated at 5%, but with a range from 5% to 37%. Apomixis, parent survival, and offspring survival were all estimated to be high in haploids, triploids, and tetraploids, and low in diploids and tetraploids (Appendix S1: Figure S6). Notably, triploid offspring survival had a bimodal peak with best values of either 5–10% or 100%; triploid parent survival had a unimodal peak with best values of 95–100%.

These predictions all should be interpreted with substantial caution because they assume all other predictors are held at median values and therefore do not satisfy normalization constraints within each row of A , nor are they conditioned on presumably more realistic values of certain parameters. Multiple parameter combinations are consistent with observations, so additional unavailable empirical data would be needed to better constrain these inferences.

DISCUSSION

The model developed in this study indicates that there are multiple feasible routes to high prevalence of odd-numbered ploidies. In the case of quaking aspen, a primarily diploid/triploid population was predicted for a wide range of parameter values and combinations. This indicates that high triploid prevalence is not actually unlikely; many processes and parameter values can yield it.

Revisiting hypotheses for aspen

Six of the 11 hypotheses put forward for the high frequency of triploids in quaking aspen were testable using this study's

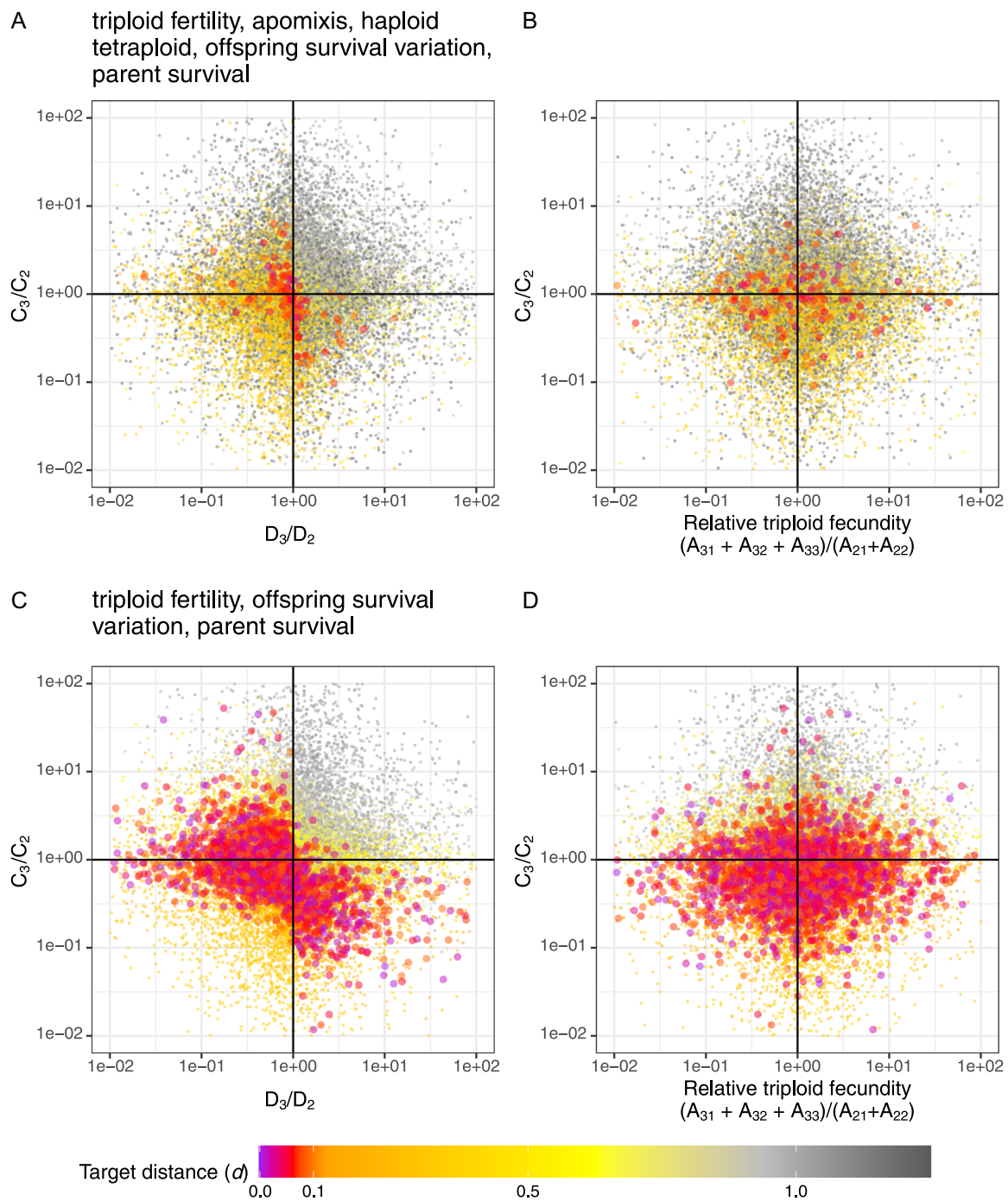


FIGURE 6 The effect of selected parameter combinations on target distance (d). Each dot represents an independent parameter sample and is colored by target distance; in-target samples yielding $d < 0.1$ are shown with larger size. The $x = 1$ and $y = 1$ lines are shown in black in all panels. In the full (triploid fertility, apomixis, haploid tetraploid, offspring survival variation, parent survival) scenario, (A) the ratio between triploid and diploid offspring survival (C_3/C_2) and the ratio between triploid and diploid parent survival (D_3/D_2) is the primary trade-off determining in-target outcomes. Note that a few parameter draws outside the plotted range are not shown for clarity and do not affect interpretation. (B) In this scenario, triploid fertility (here, relative triploid fecundity compared to diploids) is not a key driver of in-target outcomes, nor is there a trade-off observed between this fecundity and relative triploid offspring survival. In a scenario with no apomixis or haploid tetraploid life cycles, (C) the above offspring/parent survival trade-off becomes clearer, and (D) the lack of a fecundity/survival trade-off becomes clearer.

model (Table 2). Hypothesis 1 (similar fitness of diploids and triploids) was assessed as potentially plausible, because high apomixis can compensate for low fecundity; however, there was no evidence for a direct survival–fecundity

trade-off. Nevertheless, effective fecundity via parent survival (see Hypothesis 5) did increase triploid fitness. Hypothesis 2 (triploids existing as a “sink” population through high production of reduced/unreduced gametes by

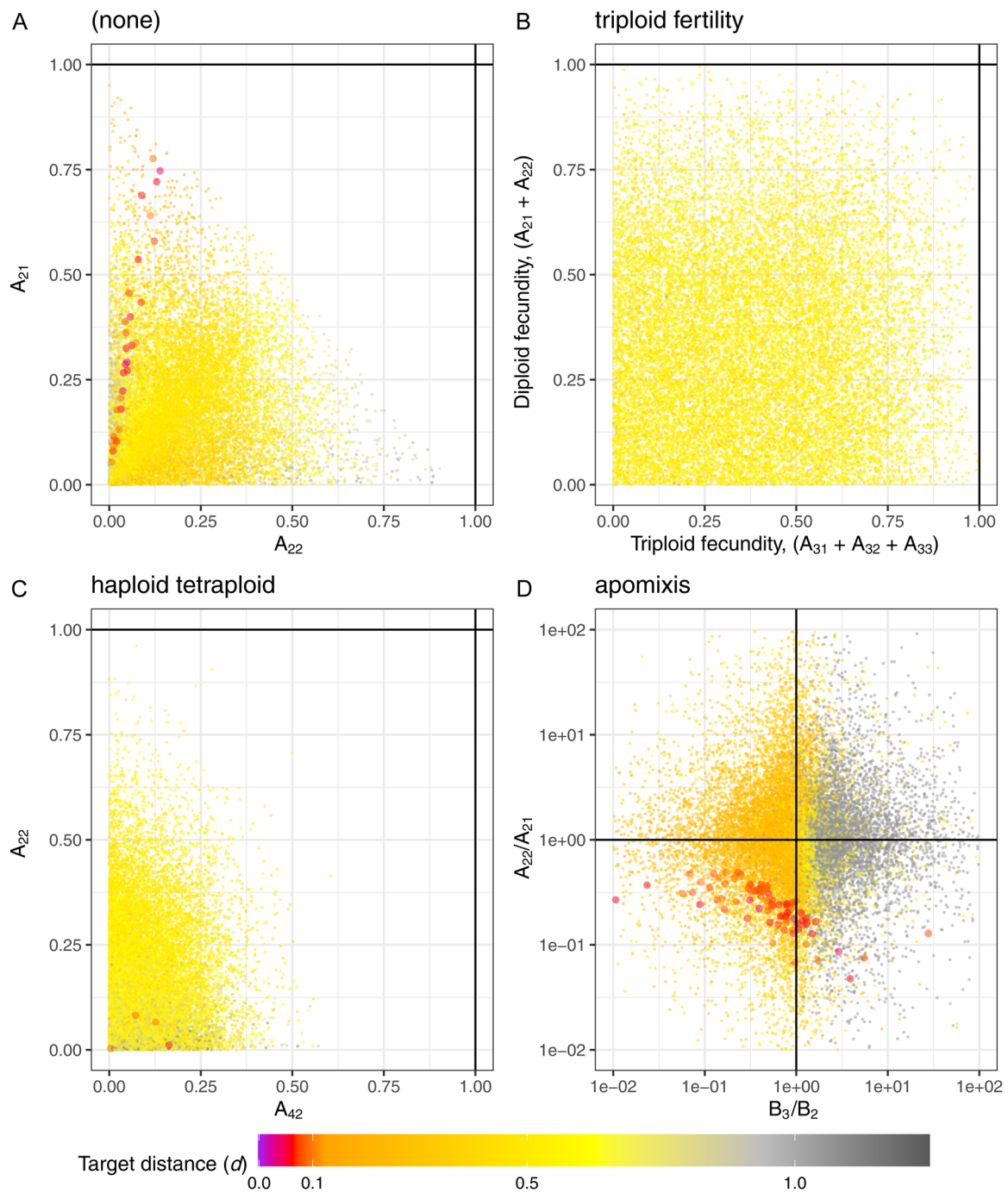


FIGURE 7 Additional scenarios reveal further parameter relationships. The effect of selected parameter combinations on target distance (d). Each dot represents an independent parameter sample and is colored by target distance; in-target samples yielding $d < 0.1$ are shown with larger size and purple-red coloration. The $x = 1$ and $y = 1$ lines are shown in black in all panels. (A) In the base scenario where only gamete production rates vary, there is a narrow band of ratios between reduced and unreduced gamete production for diploids that yields in-target outcomes. (B) In a scenario that includes triploid fertility, a range of triploid fecundities (including low/zero values) yield in-target outcomes. (C) In a scenario that also includes haploid and tetraploid cytotypes, there is a weak or unclear relationship between unreduced diploid gamete production and reduced tetraploid gamete production. (D) In a scenario that includes apomixis, there are complex trade-offs between relative triploid apomixis probabilities and unreduced gamete production by diploids.

diploids) was assessed as plausible because the base scenario could yield in-target outcomes without triploid fertility. Hypothesis 3 (triploid bridge) was assessed as rejected because the target outcome did not include tetraploids, which would be a necessary consequence of this bridge. Hypothesis 4

(apomixis) was assessed as potentially rejected. The parameter was somewhat important in the model, but the one study on the topic rejected the existence of the phenomenon in aspen (Nagaraj, 1952). More empirical investigation of apomixis in aspen is needed. Hypothesis 5 (clonality) was assessed as

TABLE 2 Hypotheses for why triploids are common. Each hypothesis is tied to a subset of model parameters; after testing, each hypothesis also yields a particular inference. The set of hypotheses is more comprehensive than could be tested with this model.

Hypothesis	Parameters varied in this study	Inference from this study
1. Odd-numbered ploidy individuals have similar fitness to even-numbered ploidy individuals via trade-offs between fecundity and growth or survival	A, \vec{C}, \vec{D}	Potentially plausible
2. Variation in production rate of gametes of each ploidy among parent ploidies	A	Plausible
3. Triploid bridge involving production of higher-fitness tetraploids that requires and maintains lower-fitness triploids	A, \vec{C}, \vec{D}	Rejected
4. Apomixis of odd-numbered ploidy parents	\vec{B}	Potentially rejected, but more empirical evidence needed
5. Vegetative reproduction (clonal growth) of odd-numbered ploidy individuals	\vec{B}, \vec{D}	Plausible
6. Ecological drift via equal fitness of ploidies	M	Potentially rejected, but more complex model may be needed
7. Mass-effect dispersal of odd-numbered ploidy individuals or gametes from a metapopulation	—	Not tested
8. Competition–colonization trade-off where metapopulation patches are preferentially colonized by odd-numbered ploidies but later outcompeted by even-numbered ploidies	—	Not tested
9. Storage effect or relative nonlinearity of competition in which temporally fluctuating environments enable coexistence of odd-numbered ploidies if they sometimes have higher fitness	—	Not tested
10. Lagged responses to historical climates in which odd-numbered ploidy individuals previously had higher fitness	—	Not tested
11. Intraspecific parasitism of odd-numbered cytotypes via resource preemption	—	Not tested

plausible because vegetative reproduction (as modeled via apomixis or parent survival) has a positive effect in the model and is consistent with field observations. Hypothesis 6 (drift) was potentially rejected because population size (M) had negligible importance in any scenario. However, this inference was tempered by the limited inclusion of density-dependent processes in the model (see later section Limitations of the model).

Thus, even if triploids are infertile (no gamete production), they can still occur at high frequencies. They can also occur even in the absence of a high frequency of tetraploids (i.e., as part of a “triploid bridge”). The primary mechanisms maintaining triploids is higher survival of triploids, either through parent or offspring survival, or via spontaneous production from unreduced and reduced gametes of diploid parents, or both. Even if triploid fertility is very low, triploids can be maintained in the population through apomixis, high parent survival, or high offspring survival, which also yields more time over which lower fertility may generate offspring. Clonality promotes all these effects. Similar ideas were previously reviewed by Herben et al. (2017).

The underlying source of triploids in the model is fusion of reduced and unreduced gametes from a diploid parent. Routes involving reduced gametes from a tetraploid parent seem unlikely because such parents were not observed, and routes involving triploid fecundity are unnecessary but helpful.

In the actual species, the true values of the underlying parameters (or whether they are even non-zero) remain unknown. The main value of this model is in identifying parameters of interest, e.g., differential survival of offspring by cytotype, differential survival of parents by cytotype, apomixis of certain cytotypes, and rates and types of gamete production by cytotype. Efforts to estimate most likely parameter values based on model optimization are likely not useful (e.g., Appendix S1: Figure S6) until a more complex spatially and temporally explicit model (e.g., including Hypotheses 7–11) is developed; in such a case, the high dimensionality of the parameter space might anyway prevent useful inference of the most-likely parameter values.

Hypotheses 1 and 5 were most consistent with the limited empirical data available. Differential survival of diploids and triploids parents is very likely based on forest

resurvey data, though the direction of the effect remains debated and likely environmentally dependent (Dixon and DeWald, 2015; Blonder et al., 2021). Differential offspring survival of diploid and triploid embryos is poorly known but potentially consistent with experimental data from (Johnsson, 1940). Recent literature seems to suggest that in hot and dry conditions, triploids have reduced survival (Blonder et al., 2021), and potentially co-occurring stand-replacing fires could then favor seed production (DeByle and Winokur, 1985; Fairweather et al., 2014) and diploid establishment. Triploid fertility seems to be low but non-zero (Goessen et al., 2022) with unreduced gametes also being produced at low frequencies (Winton and Einspahr, 1970). Tetraploids can be experimentally produced, though their survival is apparently very low (Winton and Einspahr, 1970). Other parameters, and those that would be required to test Hypotheses 7–11, remain unknown.

A key question is whether the high frequencies of diploids and triploids, and the low or zero frequencies of haploids and tetraploids are unique to aspen and, if so, why. Ramsey and Schemske (1998) reviewed frequencies of different pathways to triploidy and find strong triploid blocks in most species, but also substantial triploid fertility, suggesting that triploidy is typically but not always uncommon. Supporting this perspective, Kolář et al. (2017) found odd-numbered ploidies comprised only 11% of individuals across a 59-species study, suggesting that quaking aspen is an atypical case. However, Kolář et al. (2017) also found that odd-numbered ploidies were more common (39% of individuals) among species where asexual reproduction was more prevalent than sexual reproduction. They also highlighted the importance of assortative mating or spatial segregation of odd-numbered ploidies in their maintenance. Thus, aspen does not appear unique after accounting for its clonal nature and vegetative reproduction. Indeed, clonality may be the key property of aspen that supports its high levels of triploidy, consistent with Hypothesis 5.

Literature context

These results are consistent with the model of Van Drunen and Friedman (2022). They found that autopolyploid establishment is facilitated by clonality, equivalent to either apomixis or high parental lifespans in this model. The model of Yamauchi et al. (2004) yields somewhat similar conclusions to this study. They found that parthenogenesis of triploids promotes triploid coexistence and can overcome the triploid block (Köhler et al., 2010). They also found that tetraploid fitness must be low or else tetraploids would dominate the population. The results of the present study extend their findings with a more detailed computational exploration of the parameter space, a focus on aspen, and a conceptual exploration of additional population-level coexistence mechanisms. However, their general conclusions on evolution remain relevant and were explored in greater depth there than here. The forthcoming model of

Gaynor et al. (2023 [preprint]) also explores dynamics of mixed-ploidy populations under more spatiotemporally complex environments, but is focused primarily on tetraploids. The present results also extend the work of Husband (2004) on the role of triploids as “triploid bridges” to tetraploidy by focusing on tetraploid-free scenarios where triploids remain common. It is encouraging that qualitatively similar conclusions were obtained from these independent studies.

Limitations of the model

The model developed here has some key limitations that constrain its interpretation or use in other systems. The most obvious is the current implementation with four cytotypes, but extension to more is readily possible due to the vectorization of the underlying code. More important are the key processes left out of the model, many of which are relevant to untested Hypotheses 7–11.

First, the model necessarily elides many aspects of the species' life history, e.g., seed production, initial establishment, and growth. These factors were intentionally lumped into aggregate parameter values because they have similar effects on the phenomena of interest and are difficult to individually measure, but separating them might provide a more nuanced view of the problem. Second, the model does not consider sex, e.g., the possibility of unbalanced sex ratios and limited opportunities for sexual reproduction or interactions between cytotype and sex that may modulate parameter values. Quaking aspen is dioecious, with spatial variation in sex ratios (Mennel, 1957; Pauley and Mennel, 1957; Einspahr, 1960). Males also seem to be more common among triploids (R. Goessen, Université Laval, personal communication). However, the impacts of sex on processes considered in this model is wholly unknown, though sex-specific stress responses have been documented in the genus (Melnikova et al., 2017). Third, the model does not consider dispersal within metapopulations, excluding the possibility of processes dependent on a balance between within- and between-patch effects. Fourth, the model does not include environmental dependence for some parameters or environmental variation over time. Both effects could promote coexistence of cytotypes, e.g., via temporal niche partitioning or a storage effect (Chesson, 2000), but were not included. Along environmental gradients, many species exhibit variation in clonality (Zhang et al., 2023) and cytotype frequencies (McAllister et al., 2015) that presumably would affect the parameters of this model. Fifth, the model does not account for autodiploidization processes (Li et al., 2021) in which haploids or polyploids may spontaneously revert to diploids. All these processes could be important, but data and interpretation considerations precluded implementation in the current model version. The model of Gaynor et al. (2023 [preprint]) may be able to address some of these processes. Sixth, processes affecting the absolute size of the population are not considered,

making it difficult to fully assess the effect of drift. There is some negative density dependence implicit in the model (i.e., when one cytotype becomes less common, another necessarily becomes more common), but no explicit consideration of density-dependent processes like an Allee effect or minority cytotype advantage. Such processes may influence the coexistence of cytotypes and modulate the importance of drift.

Additionally, some concern exists about the parameter space sampling. The high dimensionality of the parameter space and the potential for multiple optima in the model response meant that exhaustive sampling or naïve application of optimization methods were not feasible. However, the limited sampling used here is not guaranteed to be sufficient to fully explore the parameter space. Test runs of individual scenarios at other sample sizes (not shown) indicated that qualitative results did not change as sample size increased. This sampling issue is therefore likely unimportant, but it is always possible that unsampled parameter combinations could substantially shift the final inferences made.

Direct validation of the model's time series predictions likely remains out of reach. Data for the dynamics of mixed ploidy populations are largely absent, except for the data of Mráz et al. (2022). In aspen, sedimentary records of pollen abundance might be used to distinguish diploid and triploid frequencies, given the larger size of triploid pollen grains (Winton and Einspahr, 1970; Every and Wiens, 1971), but complex taphonomic effects would need to be resolved.

Complications of clonality

Cytotype frequencies and fitness are typically measured assuming that all individuals are approximately interchangeable. However, in clonal species like quaking aspen, some individuals (clones) may be substantially larger than others. Because triploids tend to form larger clones, estimates of triploid frequencies may differ depending on the calculation basis. Estimation is challenging because field surveys of aspen cytotypes must be calculated either on a per-stem (per-area) basis or a per-genotype basis. Triploid clones tend to be larger in this species, which in turn can affect the estimates of target cytotype frequency and thus the estimation of in-target outcomes in this model. It is unclear what calculation basis is most appropriate. Per-genotype calculations may best reflect underlying evolutionary and population dynamics; however, per-area measurements may reflect actual landscape frequencies better and also more accurately represent intraspecific competition (e.g., due to resource preemption and priority effects from large individuals).

There are also additional effects of clonality that may be relevant to aspen that are not included in the model. First, there are potential priority effects (Fukami, 2015) relevant to triploids. Triploid clones that establish early and become large (e.g., if dispersal ability, establishment probability, or initial growth rate is higher) could become locally common.

They would then have locally dominant effects on both seed production and vegetative reproduction. Second, large clones occupy space and may disrupt pollen and seed dispersal of other genotypes in spatially complex ways, influencing the strength of intra- vs intercytotype competition. Third, there is a possibility that clones that reach a certain size become “immortal” (Peñuelas and Munné-Bosch, 2010; Klimešová et al., 2015) and drive age dependence in parameter values not included in the model. For example, some aspen clones are estimated to have ages of 10,000–70,000 years based on a molecular clock approach, though advanced age also comes with increased somatic mutations that substantially depress fertility (Ally et al., 2010). These age ranges are long relative to the timescale of interglacial periods and mean that even very rare reproduction events (difficult to capture in a model or in any contemporary field data set) may still yield high fitness for these individuals.

CONCLUSIONS

Many model parameter combinations were consistent with triploidy being common in quaking aspen, indicating it is not an unusual outcome. Despite common belief, triploid quaking aspen is probably not infertile. Even low rates of triploid fertility combined with higher triploid parent or offspring survival, yield coexistence of diploids and triploids; and even if fertility were zero, unreduced gamete production by diploids could still yield large numbers of triploids. These mechanisms do not require the existence of a “triploid bridge” to tetraploidy. Additional complex mechanisms promoting diploid–triploid coexistence could also occur, but were not investigable with this model, or with the paucity of available empirical data.

AUTHOR CONTRIBUTIONS

This contribution is by one author.

ACKNOWLEDGMENTS

I was supported by USDA NIFA grant 2022-67019-36366. I am grateful to members of the Polyploidy Webinar for feedback. I am also grateful to Alyssa Phillips, Courtenay Ray, Erin Carroll, Karen Mock, Mickey Boakye, Roosje Goessen, Roxy Cruz, and three anonymous reviewers for their feedback.

OPEN RESEARCH BADGES



This article has earned Open Data and Open Materials badges. Data and materials are available at https://github.com/bblonder/aspen_cytotype_model.

DATA AVAILABILITY STATEMENT

All code to replicate these analyses and all computed results are archived on Zenodo at <https://zenodo.org/records/10936366>.

and available on Github at https://github.com/bblonder/aspn_cytotype_model. Raw time-series outputs are provided in ZIP format to reduce archive space usage.

ORCID

Benjamin Wong Blonder  <http://orcid.org/0000-0002-5061-2385>

REFERENCES

- Ally, D., K. Ritland, and S. P. Otto. 2010. Aging in a long-lived clonal tree. *PLoS Biology* 8: e1000454.
- Anamthawat-Jónsson, K., and A. Thór Thórsson. 2003. Natural hybridisation in birch: triploid hybrids between *Betula nana* and *B. pubescens*. *Plant Cell, Tissue and Organ Culture* 75: 99–107.
- Asker, S., and L. Jerling. 1992. Apomixis in plants. CRC Press, Boca Raton, FL, USA.
- Benson, M. K., and D. W. Einspahr. 1967. Early growth of diploid, triploid and triploid hybrid aspen. *Forest Science* 13: 150–155.
- Blonder, B., C. Ray, J. Walton, M. Castaneda, K. D. Chadwick, M. Clyne, P. Gauzere, et al. 2021. Cytotype and genotype predict mortality and recruitment in Colorado quaking aspen (*Populus tremuloides*). *Ecological Applications* 31: e02438.
- Breiman, L. 2001. Random forests. *Machine Learning* 45: 5–32.
- Burton, T. L., and B. C. Husband. 2001. Fecundity and offspring ploidy in matings among diploid, triploid and tetraploid *Chamerion angustifolium* (Onagraceae): consequences for tetraploid establishment. *Heredity* 87: 573–582.
- Chesson, P. 2000. Mechanisms of maintenance of species diversity. *Annual Review of Ecology and Systematics* 31: 343–366.
- DeByle, N. V., and R. P. Winokur. 1985. Aspen: ecology and management in the western United States. USDA Forest Service General Technical Report RM-119, Rocky Mountain Forest and Range Experiment Station, Fort Collins, CO, USA.
- DeRose, R. J., K. E. Mock, and J. N. Long. 2015. Cytotype differences in radial increment provide novel insight into aspen reproductive ecology and stand dynamics. *Canadian Journal of Forest Research* 45: 1–8.
- DeWoody, J., C. A. Rowe, V. D. Hipkins, and K. E. Mock. 2008. “Pando” lives: molecular genetic evidence of a giant aspen clone in central Utah. *Western North American Naturalist* 68: 493–497.
- Dixon, G. B., and L. E. DeWald. 2015. Microsatellite survey reveals possible link between triploidy and mortality of quaking aspen in Kaibab National Forest, Arizona. *Canadian Journal of Forest Research* 45: 1369–1375.
- Einspahr, D. W. 1960. Sex ratio in quaking aspen and possible sex-related characteristics. Proceedings of the Fifth World Forestry Congress, vol. 2, 747–750, 1960, Seattle, WA, USA. USDA,
- Einspahr, D. W., J. Van Buijtenen, and J. Peckham. 1963. Natural variation and heritability in triploid aspen. *Silvae Genetica* 12: 51–58.
- Einspahr, D. W., and L. L. Winton. 1977. Genetics of quaking aspen. Research Paper WO-25. U.S. Department of Agriculture, U.S. Forest Service, Washington, D.C., USA.
- Every, A. D., and D. Wiens. 1971. Triploidy in Utah aspen. *Madroño* 21: 138–147.
- Fairweather, M. L., E. A. Rokala, and K. E. Mock. 2014. Aspen seedling establishment and growth after wildfire in Central Arizona: an instructive case history. *Forest Science* 60: 703–712.
- Farco, G. E., and M. Dematteis. 2014. Meiotic behavior and pollen fertility in triploid and tetraploid natural populations of *Campuloclinium macrocephalum* (Eupatorieae, Asteraceae). *Plant Systematics and Evolution* 300: 1843–1852.
- Fisher, R. A. 1930. The genetical theory of natural selection. Clarendon Press, Oxford, UK.
- Fox, D. T., D. E. Soltis, P. S. Soltis, T.-L. Ashman, and Y. Van de Peer. 2020. Polyploidy: a biological force from cells to ecosystems. *Trends in Cell Biology* 30: 688–694.
- Fukami, T. 2015. Historical contingency in community assembly: integrating niches, species pools, and priority effects. *Annual Review of Ecology, Evolution, and Systematics* 46: 1–23.
- Gaynor, M. L., N. Kortessis, D. E. Soltis, P. S. Soltis, and J. M. Ponciano. 2023. Dynamics of mixed-ploidy populations under demographic and environmental stochasticities. *bioRxiv* 2023.03.29.534764 [preprint]. <https://www.biorxiv.org/content/10.1101/2023.03.29.534764v1.abstract>
- Goessen, R., N. Isabel, C. Wehenkel, N. Pavy, L. Tischenko, L. Touchette, I. Giguère, et al. 2022. Coping with environmental constraints: geographically divergent adaptive evolution and germination plasticity in the transcontinental *Populus tremuloides*. *Plants, People, Planet* 4: 638–654.
- Greenwell, B. M. 2017. pdp: An R package for constructing partial dependence plots. *R Journal* 9: 421.
- Greer, B. T., C. Still, G. L. Cullinan, J. R. Brooks, and F. C. Meinzer. 2017. Polyploidy influences plant–environment interactions in quaking aspen (*Populus tremuloides* Michx.). *Tree Physiology* 38: 630–640.
- Henry, I. M., B. P. Dilkes, K. Young, B. Watson, H. Wu, and L. Comai. 2005. Aneuploidy and genetic variation in the *Arabidopsis thaliana* triploid response. *Genetics* 170: 1979–1988.
- Herben, T., J. Suda, and J. Klimešová. 2017. Polyploid species rely on vegetative reproduction more than diploids: a re-examination of the old hypothesis. *Annals of Botany* 120: 341–349.
- Hou, J., N. Ye, D. Zhang, Y. Chen, L. Fang, X. Dai, and T. Yin. 2015. Different autosomes evolved into sex chromosomes in the sister genera of *Salix* and *Populus*. *Scientific Reports* 5: 9076.
- 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.
- Jackson, R. C. 1976. Evolution and systematic significance of polyploidy. *Annual Review of Ecology and Systematics* 7: 209–234.
- Johnsson, H. 1940. Cytological studies of diploid and triploid *Populus tremula* and of crosses between them. *Hereditas* 26: 321–352.
- Kawecki, T. J. 2004. Ecological and evolutionary consequences of source-sink population dynamics. In I. A. Hanski and O. E. Gaggiotti [eds.], *Ecology, genetics and evolution of metapopulations*, 387–414. Academic Press, Oxford, UK.
- Keeler, K. H. 1990. Distribution of polyploid variation in big bluestem (*Andropogon gerardii*, Poaceae) across the tallgrass prairie region. *Genome* 33: 95–100.
- Keeler, K. H. 1998. Population biology of intraspecific polyploidy in grasses. *Population biology of grasses*, 183–206. Cambridge University Press, Cambridge, UK.
- Klimešová, J., M. P. Nobis, and T. Herben. 2015. Senescence, ageing and death of the whole plant: morphological prerequisites and constraints of plant immortality. *New Phytologist* 206: 14–18.
- Köhler, C., O. Mittelsten Scheid, and A. Erilova. 2010. The impact of the triploid block on the origin and evolution of polyploid plants. *Trends in Genetics* 26: 142–148.
- Kolář, F., M. Čertner, J. Suda, P. Schönswetter, and B. C. Husband. 2017. Mixed-ploidy species: progress and opportunities in polyploid research. *Trends in Plant Science* 22: 1041–1055.
- Levin, D. A. 1975. Minority cytotype exclusion in local plant populations. *Taxon* 24: 35–43.
- Li, Z., M. T. McKibben, G. S. Finch, P. D. Blischak, B. L. Sutherland, and M. S. Barker. 2021. Patterns and processes of diploidization in land plants. *Annual Review of Plant Biology* 72: 387–410.
- Lin, S.-J., M. Kato, and K. Iwatsuki. 1992. Diploid and triploid offspring of triploid agamosporous fern *Dryopteris pacifica*. *Shokubutsu-gakuzasshi* 105: 443–452.
- Marks, G. E. 1966. The origin and significance of intraspecific polyploidy: experimental evidence from *Solanum chacoense*. *Evolution* 20: 552–557.
- McAllister, C., R. Blaine, P. Kron, B. Bennett, H. Garrett, J. Kidson, B. Matzenbacher, et al. 2015. Environmental correlates of cytotype distribution in *Andropogon gerardii* (Poaceae). *American Journal of Botany* 102: 92–102.

- Melnikova, N. V., E. V. Borkhert, A. V. Snezhkina, A. V. Kudryavtseva, and A. A. Dmitriev. 2017. Sex-specific response to stress in *Populus*. *Frontiers in Plant Science* 8: 1827.
- Mennel, G. F. 1957. Sex ratio, hermaphroditism, and flower abnormalities in quaking aspen. *Journal of the Minnesota Academy of Science* 25: 22–25.
- Mitton, J. B., and M. C. Grant. 1996. Genetic variation and the natural history of quaking aspen. *BioScience* 46: 25–31.
- Mock, K. E., C. M. Callahan, M. N. Islam-Faridi, J. D. Shaw, H. S. Rai, S. C. Sanderson, C. A. Rowe, et al. 2012. Widespread triploidy in western North American aspen (*Populus tremuloides*). *PLoS One* 7: e48406.
- Mock, K. E., C. Rowe, M. B. Hooten, J. Dewoody, and V. Hipkins. 2008. Clonal dynamics in western North American aspen (*Populus tremuloides*). *Molecular Ecology* 17: 4827–4844.
- Mráz, P., S. Španiel, K. Skokanová, and B. Šingliarová. 2022. Temporal stability of spatial cytotype structure in mixed-ploidy populations of *Centaurea stoebe*. *AoB Plants* 14: plac052.
- Nagaraj, M. 1952. Floral morphology of *Populus deltoides* and *P. tremuloides*. Contributions from the Hull Botanical Laboratory 639. *Botanical Gazette* 114: 222–243.
- Namikawa, S., and J. Kawakami. 1934. On the occurrence of the haploid, triploid and tetraploid plants in twin seedlings of common wheat. *Proceedings of the Imperial Academy* 10: 668–671.
- Otto, S. P. 2007. The evolutionary consequences of polyploidy. *Cell* 131: 452–462.
- Pauley, S. S., and G. F. Mennel. 1957. Sex ratio and hermaphroditism in a natural population of quaking aspen. *Minnesota Forestry Notes* 55: 1–2.
- Peñuelas, J., and S. Munné-Bosch. 2010. Potentially immortal? *New Phytologist* 187: 564–567.
- R Core Team. 2022. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Website: <https://www.r-project.org>
- Ramsey, J., and D. W. Schemske. 1998. Pathways, mechanisms, and rates of polyploid formation in flowering plants. *Annual Review of Ecology and Systematics* 29: 467–501.
- Rogers, P. C., S. M. Landhäusser, B. D. Pinno, and R. J. Ryel. 2014. A functional framework for improved management of western North American aspen (*Populus tremuloides* Michx.). *Forest Science* 60: 345–359.
- Schlaepfer, D. R., P. J. Edwards, J. C. Semple, and R. Billeter. 2008. Cytogeography of *Solidago gigantea* (Asteraceae) and its invasive ploidy level. *Journal of Biogeography* 35: 2119–2127.
- Schranz, M. E., C. Dobeš, M. A. Koch, and T. Mitchell-Olds. 2005. Sexual reproduction, hybridization, apomixis, and polyploidization in the genus *Boechera* (Brassicaceae). *American Journal of Botany* 92: 1797–1810.
- Šmarda, P., L. Horová, O. Knápek, H. Dieck, M. Dieck, K. Ražná, P. Hrubík, et al. 2018. Multiple haploids, triploids, and tetraploids found in modern-day “living fossil” *Ginkgo biloba*. *Horticulture Research* 5: 55.
- Van Buijtenen, J., P. N. Joranson, and D. W. Einspahr. 1957. Naturally occurring triploid quaking aspen in the United States. Proceedings of the Society of American Foresters, 62–64, 1957, Syracuse, NY, USA.
- Van de Peer, Y., T.-L. Ashman, P. S. Soltis, and D. E. Soltis. 2021. Polyploidy: an evolutionary and ecological force in stressful times. *Plant Cell* 33: 11–26.
- Van Drunen, W. E., and J. Friedman. 2022. Autopolyploid establishment depends on life-history strategy and the mating outcomes of clonal architecture. *Evolution* 76: 1953–1970.
- Wickham, H., M. Averick, J. Bryan, W. Chang, L. D. McGowan, R. François, G. Grolemund, et al. 2019. Welcome to the tidyverse. *Journal of Open Source Software* 4: 1686.
- Winton, L., and D. W. Einspahr. 1970. Tetraploid aspen production using unreduced triploid pollen. *Forest Science* 16: 180–182.
- Wright, M. N., and A. Ziegler. 2017. ranger: A fast implementation of random forests for high dimensional data in C++ and R. *Journal of Statistical Software* 77: 1–17.
- Yamauchi, A., A. Hosokawa, H. Nagata, and M. Shimoda. 2004. Triploid bridge and role of parthenogenesis in the evolution of autopolyploidy. *American Naturalist* 164: 101–112.
- Zhang, H., S.-C. Chen, S. P. Bonser, T. Hitchcock, and A. T. Moles. 2023. Factors that shape large-scale gradients in clonality. *Journal of Biogeography* 50: 827–837.
- Zielinski, M.-L., and O. Mittelsten Scheid. 2012. Meiosis in polyploid plants. In P. S. Soltis and D. E. Soltis [eds.], *Polyploidy and genome evolution*, 33–55. Springer-Verlag, Berlin, Germany.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Blonder, B. W. 2024. Why are triploid quaking aspen (*Populus tremuloides*) common? *American Journal of Botany* e16325. <https://doi.org/10.1002/ajb2.16325>