

This is a PDF file of an article that is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain. The final authenticated version is available online at: <https://doi.org/10.1073/pnas.2400018121>

For the purpose of Open Access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

This work was funded by European Research Council (DOUBLE-TROUBLE 833522)

Interspecific transfer of genetic information through polyploid bridges

Felipe Kauai^{a,b,c,1}, Quinten Bafort^{a,b}, Frederik Mortier^{a,b,c}, Marc Van Montagu^{a,b,1}, Dries Bonte^{c,1}, Yves Van de Peer^{a,b,d,e,1}

ORCID ID: 0000-0002-4991-8256 (F.K.), 0000-0003-2155-3344 (Q.B.), 0000-0003-4565-5038 (F.M.), 0000-0003-4711-5131 (M.V.M.), 0000-0002-3320-7505 (D.B.), and 0000-0003-4327-3730 (Y.V.d.P.)

^aDepartment of Plant Biotechnology and Bioinformatics, Ghent University, 9052 Ghent, Belgium; ^bCenter for Plant Systems Biology, VIB, 9052 Ghent, Belgium; ^cDepartment of Biology, Terrestrial Ecology Unit, Ghent University, 9000 Ghent, Belgium; ^dCentre for Microbial Ecology and Genomics, Department of Biochemistry, Genetics and Microbiology, University of Pretoria, Pretoria 0028, South Africa; ^eCollege of Horticulture, Academy for Advanced Interdisciplinary Studies, Nanjing Agricultural University, Nanjing 210095, China

¹To whom correspondence may be addressed. Email: felipekauai.pereira@psb.vib-ugent.be; marc.vanmontagu@ugent.be, dries.bonte@ugent.be, or yves.vandeppeer@psb-vib.ugent.be.

Author contributions: F.K., Q.B., F.M., M.V.M., D.B., and Y.V.d.P. designed research; F.K., Q.B., and F.M. performed research; F.K., Q.B., F.M., D.B., Y.V.d.P analyzed data; All authors contributed to the writing of the original and revised manuscript.

The authors declare no conflict of interest.

Abstract

Hybridization blurs species boundaries and leads to intertwined lineages resulting in reticulate evolution. Polyploidy, the outcome of whole genome duplication (WGD), has more recently been implicated in promoting and facilitating hybridization between polyploid species, potentially leading to adaptive introgression. However, because polyploid lineages are usually ephemeral states in the evolutionary history of life it is unclear whether WGD-potentiated hybridization has any appreciable effect on their diploid counterparts. Here, we develop a model of cytotype dynamics within mixed-ploidy populations to demonstrate that polyploidy can in fact serve as a bridge for gene flow between diploid lineages, where introgression is fully or partially hampered by the species barrier. Polyploid bridges emerge in the presence of triploid organisms, which despite critically low levels of fitness, can still allow the transfer of alleles between diploid states of independently evolving mixed-ploidy species. Notably, while marked genetic divergence prevents polyploid-mediated interspecific gene flow, we show that increased recombination rates can offset these evolutionary constraints, allowing a more efficient sorting of alleles at higher-ploidy levels before introgression into diploid gene pools. Additionally, we derive an analytical approximation for the rate of gene flow at the tetraploid level necessary to supersede introgression between diploids with non-zero introgression rates, which is especially relevant for plant species complexes, where interspecific gene flow is ubiquitous. Altogether, our results illustrate the potential impact of polyploid bridges on the (re)distribution of genetic material across ecological communities during evolution, representing a potential force behind reticulation.

Keywords: whole-genome duplication, polyploidy, introgression, reticulate evolution

Significance statement

Polyploidy, which results from the duplication of an organism's entire genome, is a known source of genetic novelty that promotes speciation, thus contributing to the evolution and distribution of biodiversity. Its eco-evolutionary interactions with coexisting diploid relatives are nonetheless elusive. Here, we have developed a theoretical model to understand the temporal dynamics of mixed-ploidy populations, revealing that polyploidy may serve as a bridge for gene flow across diploid species boundaries. Due to exceptional hybridization potential and the intricate mating dynamics specific to mixed-ploidy populations, polyploids may play an important role in the emergence of intertwined lineages. Our findings provide, therefore, insights into the mechanisms that underlie the distribution of genetic material within ecosystems, leading to reticulate evolutionary histories.

Introduction

The exchange of genetic information across species barriers leads to complex interconnections between lineages and promotes reticulated evolutionary histories. This phenomenon has been well-documented in bacteria, where horizontal gene transfer significantly impacts the spread of DNA fragments among various lineages, challenging the traditional concept of bacterial species (1). However, recent findings from the expanding collection of sequenced genomes reveal that the transfer of genetic material between different species is common across all forms of life (2-5). These instances of gene transfer between species have a significant impact on how we reconstruct evolutionary relationships (6), influencing our understanding of the diversification of life as a whole. Identifying the pathways through which genetic material moves across species is therefore essential to better understand the mechanisms underlying the emergence and maintenance of biodiversity.

Polyploidy, the outcome of whole genome duplication (WGD), or the addition of one or more complete sets of chromosomes to a diploid genome, is widely recognized to have important implications for the generation of genetic novelty over macroevolutionary scales (7-11). Interspecific gene flow among polyploids, also referred to as WGD-mediated introgression, has been more recently implicated in promoting adaptation of polyploid species to specific intracellular demands and challenging environments (12-21). Transfer of genetic material between the tetraploid cytotypes of *Arabidopsis arenosa* and *A. lyrata*, for instance, was shown to stabilize meiosis among the higher ploidy individuals of both species (15). Introgression mediated by WGD might also have conferred a selective advantage to polyploid amphibians in adapting to extreme environments (16), and is likely to be responsible for the highly reticulate evolutionary history of fungi (4, 22). In fact, artificially induced polyploidy is long known among plant breeders to facilitate hybridization of otherwise sterile diploid hybrids, allowing the generation of agriculturally useful species (23, 24). In all respects, these observations support the hypothesis that polyploidy may indeed break the species barrier (16).

WGD is regarded as an extreme mutational event, and its emergence is often linked to environmental turmoil (10, 25-29). Ongoing efforts to uncover WGDs in the history of life (30, 31) reveal that several of such events seem to cluster around periods of environmental upheaval (7, 32, 33), such as the mass extinction event at the Cretaceous-Paleogene (K-Pg) boundary (34, 35), indicating a potential enhanced capacity of polyploids to adapt and respond to stressful conditions (27, 36-39). Notwithstanding the presumed adaptive potential, phylogenetic and theoretical analyses also suggest that polyploid lineages exhibit higher extinction rates than

diploids (40-42), or eventually evolve back to the (more) stable diploid state through the process of rediploidization (33). Therefore, due to its usually ephemeral presence in the history of life, it is unclear whether polyploidy has an appreciable effect on the evolutionary trajectories of their diploid counterparts, despite its potential to breach the species barrier. As there is no clear route known for gene flow from higher-ploidy levels back to diploids, WGD-mediated adaptive introgression is expected to operate uniquely at the (transient) polyploid level (21, 43), e.g., from a diploid to its derived tetraploid towards the tetraploid of another species.

Current theory neglects the presence, even if temporary, of intermediate odd-ploidy levels. For example, triploid cytotypes, known to exhibit low overall fitness (44), may function as genetic bridges between tetraploids and their diploid counterparts through the generation of gametes compatible with both ploidy levels. The discovery that the genes coding for the C_4 photosynthetic pathway in the tetraploid *Neurachne muelleri* and the diploid, tetraploid and hexaploid cytotypes of *Neurachne munroi* are not paralogous, but likely shared through hybridization or lateral gene transfer (18) suggests that the introgression of alleles through WGD may indeed spread among cytotypes and profoundly impact the evolutionary history of the recipient species. Clear signs of admixture at the polyploid level and intraspecific ploidy variation in other grass lineages, such as *Alloteropsis* (45, 46) suggest that lateral gene transfer, mediated by WGD, does occur (47), and that genetic information can introgress into the (more) stable diploid levels. If this is the case, the duplication of the entire genome of an organism may not only promote the generation of genetic novelty, but also the distribution of extant genetic material among species, thereby catalyzing reticulation. Since polyploidy incidence in plants is strikingly common (48, 49), as well as in several animal clades (38, 50), a systematic understanding of gene flow and cytotype dynamics within and among mixed-ploidy systems is paramount.

According to the current view, polyploids primarily emerge by the union of unreduced gametes, which are formed by a rare event resulting from the non-segregation of chromosomes during meiosis (44, 51, 52). For example, unreduced (diploid) gametes may fuse within a population of diploid organisms to give rise to a new tetraploid cytotype, which then is expected to face significant challenges to establishment due to negative frequency-dependent selection (53). That is, polyploid establishment is contingent on successful mating, which depends on the frequency of other organisms displaying the same ploidy level. However, triploid intermediates are hypothesized to relieve the frequency-dependent disadvantage of the emerging tetraploid cytotype by functioning as bridges between ploidy levels (44, 54-57). This observation is supported by a number of experimental efforts, which indicate that triploids may also promote

bidirectional interploidy gene flow in many systems, despite critically low levels of fitness relative to diploids and tetraploids (43, 44, 57-65).

To understand gene flow within and between mixed-ploidy populations, here we generalize the theoretical framework for cytotype dynamics introduced by Felber (66), which models the emergence of autopolyploids, *i.e.*, polyploid organisms emerging within a single species. We provide a new analytical account of the model by including triploid cytotypes with variable fitness, which allows us to derive the equilibrium cytotype frequencies and gauge the role of triploids within mixed-ploidy populations as far as their interactions with tetraploid and diploid cytotypes are concerned. Next, we extend this model to a broader theoretical framework consisting of two independent mixed-ploidy populations that are partially connected at the tetraploid level, *i.e.*, displaying WGD-potentiated hybridization, using a stochastic individual-based simulation. We then set out to explore the conditions and constraints for WGD-mediated interspecific gene flow between diploid cytotypes of different mixed-ploidy species, and quantify how much gene flow at the tetraploid level is necessary to override signals of direct introgression between the involved diploid species. Altogether, our results provide insights into the extent to which polyploids could contribute to reticulation by WGD-mediated interspecific gene flow, despite their usually transient existence in ecological systems.

Results

Model for cytotype dynamics in a single mixed-ploidy population. We assume an initially diploid population consisting of N individuals which reproduce sexually in discrete non-overlapping generations. Specifically, we assume that in each generation an effectively infinite pool of gametes is formed, of which a random sample of gametes is obtained to generate the next generation. Gametes can either be haploid or diploid. When two haploid (diploid) gametes fuse, a diploid (tetraploid) individual is produced, whereas the fusion of haploid and diploid gametes produces a triploid individual. Diploid organisms produce diploid gametes (unreduced) with probability v and haploid gametes with probability $(1 - v)$. The emerging triploid cytotypes are assumed to display reduced fitness ϕ , due to aneuploidy or reduction in viability, for example. Triploids produce either haploid or diploid gametes, with equal probability, as studied by previous theoretical models (54, 55) (see Discussion for an account of triploid gamete formation). Tetraploids, on the other hand, are allowed to generate only diploid gametes, as there is no known mechanism of viable haploid gamete formation by tetraploid

cytotypes (43) (but see (67, 68) for examples of haploid gametes generated by tetraploid cytotypes in fish).

We define such a population, which consists of different cytotypes that emerge through the formation and random fusion of different gamete types as a mixed-ploidy population. Notice that because only haploid and diploid gametes are allowed in our model (across all ploidy levels), only three cytotypes can emerge by construction, namely, diploids, triploids, and tetraploids. The emergence of new cytotypes in the population is therefore a function of the random fusion of gamete types and is, as such, conditioned on the cytotypes in the population only. It is instructive to consider the cytotypic frequency dynamics in the limit $N \rightarrow \infty$, as studied by Felber (66), where we can describe the time evolution of cytotypic frequencies within such a mixed-ploidy population through a set of non-linear recursive equations. Let g_p be the frequency of the p -ploidy gamete in the gamete pool produced, where $p \in \{1, 2\}$. Then, after one generation of random mating one can compute the frequencies of each gamete type as follows:

$$g_1' = \frac{g_1^2(1 - v) + \varphi g_1 g_2}{g_1^2 + 2\varphi g_1 g_2 + g_2^2} \quad [1]$$

$$g_2' = \frac{v g_1^2 + \varphi g_1 g_2 + g_2^2}{g_1^2 + 2\varphi g_1 g_2 + g_2^2} \quad [2]$$

To precisely understand the formulation in Eqs 1 and 2, let x_2 , x_3 and x_4 represent the frequency of diploid, triploid and tetraploid cytotypes, respectively. In a mixed-ploidy population which produces haploid and diploid gametes with frequencies g_1 and g_2 , respectively, the expansion of the binomial $(g_1 + g_2)^2$ returns the expected frequency of each cytotypic: $x_2 = g_1^2$, $x_3 = 2g_1 g_2$ and $x_4 = g_2^2$. By reducing the contribution of triploids to the gamete pool by a factor φ , we arrive at the formulation for the dynamics of gametes in Eqs. 1 and 2. Then, gamete equilibrium frequencies can be obtained by setting $g_1' = g_1$ and $g_2' = g_2$, and solving for g_1 (note that $g_2 = 1 - g_1$). We find that the equilibrium gamete frequencies satisfy:

$$g_{1+}^*, g_{1-}^* = \frac{3\varphi + v - 3}{4(\varphi - 1)} \pm \sqrt{\frac{(3 - 3\varphi - v)^2}{16(\varphi - 1)^2} - \frac{1}{2}} \quad [3]$$

Eq. 3 tells us that two fixed points for g_1 (+ and -) exist for appropriate values of triploid fitness φ and unreduced gametes frequency v , one of which is unstable and the other one stable (see *SI Appendix Text S1*). The stable equilibrium frequencies, which the system always converges to under the assumed initial conditions (an initial population of diploids), are shown for each cytotypic as a function of v and φ (Fig. 1A). The rate at which unreduced gametes are

produced by diploids has a direct influence on the frequencies of each cytotype in the population. When $\varphi = 0$, i.e., triploids are completely unviable, we observe that tetraploids overtake the system ($x_4 = 1$) for $v = 0.171$. This is the maximum limit of unreduced gametes frequency v' numerically shown by Felber (66) to allow stable coexistence between diploids and tetraploids in the absence of triploids and external selective pressures. Our formulation of Felber's model allowed us to generalize his results to the case where triploids inhabit the system with any fitness ($\varphi > 0$), and obtain an explicit analytical expression for v' given any φ as follows:

$$v' = (\varphi - 1)(2\sqrt{2} - 3) \quad [4]$$

As the probability of unreduced gametes v increases, the frequency of triploid and tetraploid cytotypes becomes progressively higher until a value of v' is reached (see *SI Appendix Text S1* for the proof of the result in Eq. 4), over which tetraploids dominate. For instance, with triploid fitness $\varphi = 0.30$, Eq. 4 tells us that v' is 12.01%, and below this threshold cytotypes will coexist, i.e., display stable non-zero frequencies. We confirmed this expectation by numerically iterating Eqs. 1 and 2 and comparing with an individual-based stochastic simulation of the model with finite population sizes (Fig. 1B).

The influence of triploid fitness on the stable frequencies of cytotypes within a population is anticipated by the triploid bridge hypothesis, where triploids may aid the emergence of tetraploids by reducing the effect of *minority cytotype exclusion* (55). To understand it in a qualitative manner, notice the probability that a triploid is formed within a diploid population is $2(1 - v)v$, whereas the formation probability of a tetraploid is simply v^2 , with $v \ll 1$. As we increase triploid fitness, the frequencies of triploid cytotypes converge to higher magnitudes in the system. By extension, the relative proportion of diploid gametes increases as they are produced with probability v in diploids and probability 0.5 in triploids (Eq. 2). The higher load of diploid gametes within the population, essentially due to a partial replacement of diploids by triploids, then facilitates the expansion of tetraploid cytotypes. Therefore, despite reduced fitness, triploids can relieve the frequency-dependent disadvantage of tetraploids, aiding their establishment by introducing a higher number of diploid gametes in the system, in accordance with previous theory (44, 55). This observation is further supported by several experimental studies across many plant families, where intermediate (odd-ploidy) cytotypes are often found and hypothesized to have a significant impact on admixture (57). Similar patterns can also be inferred, albeit to a lesser extent, in animal taxa (see Text S2 in *SI Appendix* for a more extensive review of interploidy admixture in both plants and animal lineages).

Mixed-ploidy model of interspecific gene flow. To understand how interspecific gene flow can occur at the diploid level through polyploidization, we have built a stochastic individual-based model of two discrete finite panmictic mixed-ploidy populations. Both populations follow largely independent dynamics, as described in the previous subsection, but are connected by migration with rate κ at the tetraploid level, which represents the degree with which tetraploids break the species barrier (16, 21). The initial population consists of a set of identical individuals bearing a diploid genome represented by two finite chromosomes with L biallelic unlinked loci. At each mating event, recombination and gamete formation take place (see *Methods*), with offspring emerging from the union of its parents' gametes. Fig. 2 illustrates the logical structure of the model, whereby two mixed-ploidy species (A and B) are connected only by migration at the tetraploid level. In this work, we analyze interspecific gene flow in a system that has triploids in both mixed-ploidy populations, for conceptual clarity. However, asymmetrical gene flow is also expected, as one of the populations might lack triploid cytotypes due to restrictively low fitness. Triploid intermediates are not required for gene flow from diploids to tetraploids but are required for introgression from tetraploids back to diploids (43, 69). A lack of triploid cytotypes in one mixed-ploidy population thus allows for the emergence of incomplete bridges with asymmetrical gene flow only to the diploid in the population with triploids (see *SI Appendix* Fig. S1).

Let G_A and G_B be haploid sequences that represent the reference genome of species A and B, respectively. Then, we consider a hypothetical divergent lineage, where each mixed-ploidy population represents one of two species that are genetically or ecologically (see Discussion) isolated with genetic distance $d(G_A; G_B)$ (see Fig. 3A and *Methods* for technical details). We first assume a neutral environment, i.e., the genotype of each individual has no bearing on how well it performs within its population nor on the genetic compatibility between two reproducing organisms. The process of gene transfer between populations down to diploid levels is then controlled by two parameters: the degree to which polyploids from one population enter the gene pool of the next population, represented by κ , and triploid fitness φ . The system is initialized so that in each population diploid individuals are fixed for a different allele in each one of the L loci (see *SI Appendix Algorithm S1* for technical details on code implementation). Then, as time unfolds, migration at the tetraploid level carries genetic information from one mixed-ploidy population to the next at rate κ . We then measured the frequency of introgressed

alleles at the diploid state per generation and verified that fixation of non-native alleles can happen even for critically low values of κ (Fig. 3B).

Once a non-native tetraploid enters the new mixed-ploidy population, its genetic information can spread by three different non-exclusive pathways; a $4x \times 4x$ cross, in which case the new genetic information remains at the tetraploid level, and $4x \times 3x$ or $4x \times 2x$ cross, which may produce either a tetraploid or a triploid individual. Here, we are primarily concerned with the latter two pathways, as a triploid organism can then produce haploid gametes and backcross with the diploid state inside the population producing a new diploid cytotype. The speed and frequency with which alleles flowing through these pathways enter the diploid state of the new population will then depend on triploid fitness and random effects produced by drift (see Methods). We found that the frequency of introgressed alleles per generation as a function of φ is qualitatively similar to the effect of κ . Even for $\varphi = 0.01$, the frequency of non-native alleles grows to an average of 20.7% after 20,000 generations (Fig. 3C). It should be noted, however, that this effect is in part a consequence of the lower frequency of tetraploid cytotypes in both populations, as described by Eqs. 1 and 2. By decreasing φ and keeping v constant, the frequency of tetraploid cytotypes will decay. Therefore, in *SI Appendix Text S3* we tested the effect of triploid fitness by normalizing v as a function of φ (Eq. 4), which allowed us to show the same effect decoupled from tetraploid frequencies. Also, in *SI Appendix Text S4*, we further show the evolution of introgression using a deterministic model of a single biallelic locus in the limit $N \rightarrow \infty$, where we verify that in the absence of selective pressures, fixation of a non-native neutral allele happens in the span of a few generations.

As each population is fixed for a different allele symmetrical gene flow at the tetraploid level will eventually lead to an equilibrium frequency of 50% of allele types in each population given enough time. Therefore, we were interested in the frequency with which a single mutant allele - subject to drift (see *Methods*) - arising in say, population A, gets established in population B. We initialized our system with both populations fixed for a wild-type variant and after cytotype equilibrium frequencies have been reached, as described by Eq. 3, we released a mutant neutral allele in population A at the diploid level and measured the number of times the mutant appears in the diploid gene pool of population B or gets lost. We found the probability that the mutant survives and introgresses into the diploid state of population B to be $\sim 2.3\%$, with the time it takes to cross the bridge well described by an exponential distribution with an average of ~ 43 generations (see *SI Appendix Fig. S2*).

To further understand gene flow through polyploid bridges, we built a fitness map that modulates the probability of successful mating with respect to the genetic distance between individuals and the population's optimal genome (see Methods, Eq. **16**). Essentially, the probability that an individual successfully reproduces depends on how well its genetic background approximates an optimal sequence (reference) that represents the environment of its population. We first tested the effect that selection strength, represented by s , has on the average frequency of introgressed alleles in each population (Fig. 3D). As expected, the larger the selection strength is, the lower the frequency of non-native alleles that can enter a new gene pool at the diploid level, impeding the system from achieving equilibrium. When genetic divergence is too high, hybrids are subject to strong selective pressures, and introgression of non-native alleles is blocked (Fig. 3E). In our simulations, introgression at the diploid level under high genetic divergence between populations is primarily hampered by high selection at the tetraploid state (see *SI Appendix* Fig. S3). Higher recombination rates can, however, increase optimal allele sorting between lineages for introgression down to the diploid states (see *SI Appendix* Fig. S4).

WGD-mediated adaptive introgression compared to adaptive introgression through diploids. We now turn our attention to the transfer of genes with adaptive value into a new diploid gene pool from a different mixed-ploidy population. WGD-mediated adaptive introgression has been argued to play a substantial role in the adaptation of polyploids for both intracellular and environmental demands (21), but its relevance may extend to diploids by means of the mechanisms explored previously. Since reproductive isolation between diploid species is seldom complete (70), especially in plants (20), the question remains how much gene flow is necessary at the tetraploid level (κ) to supersede diploid-mediated adaptive introgression. If diploid states exchange genetic information interspecifically, even if at low rates, we ask how much gene flow at the tetraploid level is required to override diploid-mediated introgression of an adaptive allele into a new diploid gene pool. This question arises naturally when one asks about the relative importance of gene flow between polyploid taxa given that their diploid counterparts are not completely isolated. More specifically, we are interested in the time it takes for an allele with adaptive value to introgress into a new diploid gene pool through polyploid bridges and compare it with interspecific gene flow at the diploid level (Fig. 4).

To measure how introgression of a beneficial allele between tetraploids differs from gene flow between diploid states, we built a mainland-island model with unidirectional gene flow from the mainland (species A) into the island (species B) (71) for each ploidy-mediated introgression case. We assumed a single biallelic locus, where the mainland species is fixed for a beneficial allele A (both diploids and tetraploids) and the island species is fixed for a deleterious variant a (across all cytotypes). The deleterious allele has a negative additive fitness effect s on genotypes, whose fitness of the homozygous states is decreased by a factor $(1 - s)$. Here, we shall derive the equations that describe adaptive introgression through the route $4x$ (species A) \rightarrow $4x$ (species B) \rightarrow $3x$ (species B) \rightarrow $2x$ (species B), *i.e.*, polyploidy-mediated introgression. For the tetraploid gene pool in the island there are three possible gametes, which we define as $D_1 \triangleq AA$, $D_2 \triangleq Aa$ and $D_3 \triangleq aa$. Gametes of type D_1 migrate from the tetraploid gene pool at the mainland at a rate κ , which will mix with other diploid gametes produced within the island by tetraploid and triploid genotypes (as described in previous sections). Let p_t and r_t represent the frequency of A in the island tetraploid and triploid gene pools, respectively, at time t , and $q_t = 1 - p_t$ and $w_t = 1 - r_t$ the frequency of their corresponding deleterious variants. Under random mating and assuming no double reduction during meiosis, with average triploid fitness $\varphi = 0.30$ (44), we found that the proportions of each diploid gamete type in the island gene pool, following unidirectional gamete migration from the mainland, can be calculated as follows:

$$D_1 = \{p_t^4 + 2p_t^3q_t + p_t^2q_t^2\} \left(1 - \kappa - \frac{\varphi}{2}\right) + \kappa + \frac{1}{2}\varphi(r_t^3 + r_t^2w_t) \quad [5]$$

$$D_2 = \{2p_t^3q_t + 4p_t^2q_t^2 + 2p_tq_t^3\} \left(1 - \kappa - \frac{\varphi}{2}\right) + \frac{1}{2}\varphi(2r_t^2w_t + 2r_tw_t^2) \quad [6]$$

$$D_3 = \{p_t^2q_t^2 + 2p_tq_t^3 + q_t^4\} \left(1 - \kappa - \frac{\varphi}{2}\right) + \frac{1}{2}\varphi(r_tw_t^2 + w_t^3) \quad [7]$$

Likewise, we defined haploid gametes in the island as $H_1 \triangleq A$ and $H_2 \triangleq a$, which can be produced by triploid and diploid cytotypes, as in the previous models. Then, if g_t and $z_t = 1 - g_t$ are the frequencies of A and a , respectively, among the diploid cytotypes, the proportion of haploid gamete types, following unidirectional migration from the mainland diploid species into the island (at rate β) and random mating, can be calculated as follows:

$$H_1 = \{g_t^2 + g_tz_t\} \left(1 - \beta - \frac{\varphi}{2}\right) + \beta + \frac{1}{2}\varphi(r_t^3 + 2r_t^2w_t + r_tw_t^2) \quad [8]$$

$$H_2 = \{g_tz_t + z_t^2\} \left(1 - \beta - \frac{\varphi}{2}\right) + \frac{1}{2}\varphi(r_t^2w_t + 2r_tw_t^2 + w_t^3) \quad [9]$$

Eqs. 5 to 9 thus define the proportions of all gamete types produced by all cytotypes within the island at each time t with ongoing gene flow from the mainland. By computing the

gamete types, we can retrieve the genotypes that make up the island population simply by the expansion of the polynomial $(D_1 + D_2 + D_3 + H_1 + H_2)^2$, whose frequencies will decay according to additive selection on the deleterious variant. For example, the resulting tetraploid genotype D_2^2 will have its frequency decreased by a factor $(1 - 0.5s)$ as half of its loci are a . By collecting all terms of the expansion and computing the frequency of A within the diploid gene pool, we find that the recursive equation that describes the evolution of the beneficial mutation among the diploid cytotypes is written as follows:

$$g_{t+1} = \frac{1}{\Psi} \{H_1^2 + H_1 H_2 (1 - 0.5s)\} \quad [10]$$

where,

$$\Psi = H_1^2 + 2H_1 H_2 (1 - 0.5s) + H_2^2 (1 - s) \quad [11]$$

In *SI Appendix Text S5*, we provide a detailed derivation of Eqs. 5 to 9, prove the result in Eq. 10, and provide the explicit formulas for p_{t+1} and r_{t+1} . Notice that when $\varphi = 0$, the only source of the beneficial mutation in the diploid gene pool comes from diploid-mediated introgression at rate β (Eqs. 8 and 9). When $\beta = 0$, beneficial mutations only arrive through polyploidy-mediated gene flow (if $\varphi > 0$), and thus g_{t+1} is only a function of κ .

Setting $\varphi = 0$ and $\beta = \kappa$, we first measured how the frequency of allele A unfolds in time for introgression at the tetraploid level ($4x \rightarrow 4x, p_t$) and compared it with introgression at the diploid level ($2x \rightarrow 2x, g_t$) for different values of s . That is, for the same introgression rate between tetraploids and diploids, we ask how introgression unfolds between tetraploid and between diploid populations. As expected, the speed with which allele A reaches fixation in the receiver populations depends on how strong selection is on the deleterious variant a , that is, the higher the selection strength on a , the faster introgression occurs (Fig. 5A). Additionally, for any selection coefficient $s > 0$, diploid-mediated introgression is always more efficient (*SI Appendix Fig. S6*). Indeed, population genetics theory predicts that deleterious mutations may accumulate faster in polyploids relative to diploids due to a masking effect (72), which has been more recently experimentally confirmed by a number of studies (29, 73). Essentially, due to the higher number of genome copies in the polyploid state, deleterious mutations are masked from selection, and therefore are expected to accumulate in polyploid lineages with much less friction. Alternatively, our results show that this phenomenon decreases the efficiency with which the introgression of a beneficial allele occurs in a tetraploid as compared to a diploid gene pool.

To understand how both two routes ($4x \rightarrow 4x \rightarrow 3x \rightarrow 2x$ and $2x \rightarrow 2x$) differ for an introgressing beneficial allele, we measured the time with which the beneficial mutation gets

fixed within the diploid gene pool through diploid-mediated gene flow for low introgression rates ($\beta \ll 1$), and computed the WGD-mediated introgression rate (κ_{min}) for which the allele gets fixed in the same diploid pool within the same time interval (see *SI Appendix Algorithm S2* for a description of the algorithm). We then computed the factor $\theta = \kappa_{min}/\beta$ as it more clearly informs us on the minimum increase in WGD-mediated introgression rate necessary for introgression through polyploidy to be at least as efficient as diploid-mediated adaptive introgression. We found that θ is a linear function of the selection coefficient s (Fig. 5B), that is, for each unit increase in s the amount of WGD-mediated gene flow to override diploid-mediated introgression increases linearly. However, the degree of the increase (slope) as well as the minimum increase for a neutral allele (intercept when $s = 0$) has a complex relationship with the introgression rate at the diploid level (β). Thus, in general, we can write θ as a function of β as follows:

$$\theta(\beta) = \lambda(\beta)s + \eta(\beta) \quad [12]$$

That is, θ is a linear function of the selection coefficient s , where the slope $\lambda(\beta)$ and intercept $\eta(\beta)$ are unknown functions of β . To find these unknown functions we modelled separately the slope and intercept with respect to the introgression rate at the diploid level. Both the slope $\lambda(\beta)$ (Fig. 5C) and intercept $\eta(\beta)$ (Fig. 5D) were found to be non-trivial functions of β , whose forms we could only approximate locally ($0.001 \leq \beta \leq 0.05$). These approximations allowed us to retrieve a general equation to estimate the minimum amount of gene flow at the tetraploid level necessary to override adaptive introgression mediated by diploid states. Realizing $\kappa_{min} = \theta\beta$, we write:

$$\begin{aligned} \kappa_{min} \approx \beta \left\{ \left[\frac{0.37}{\beta^{0.91}} e^{-0.14\beta} + 0.14 \ln(\beta) \right] s \right. \\ \left. + [-22.95\sqrt{\beta}e^{-9.07\beta} + 1.06 \ln(\beta) + 8.36] \right\} \end{aligned} \quad [13]$$

For any s , Eq. 13 tells us that any $\kappa > 0$ is sufficient to allow WGD-mediated introgression when diploids are isolated ($\beta \rightarrow 0$) consistent with our arguments in the previous sections. The use of Eq. 13 is subject to two experimental challenges in natural or artificial systems. First, the introgression rate at the diploid level β must be recovered, which can be determined by analyses of gene flow using molecular markers (74), for example. A second challenge concerns the determination of the selection coefficient s that a given allele is subject to relative to an adaptive variant present in the population from which gene flow is being measured. However, a first lower-bound to gene flow at the tetraploid level necessary to override gene flow at the diploid level can be achieved by setting $s = 0$, which informs the

experimentalist on the amount of gene flow between tetraploids necessary for polyploid bridges to display a relevant impact on the distribution of neutral genetic material across species boundaries.

Discussion

Understanding how genetic material flows among species is crucial for our understanding of reticulate evolution. By generalizing the well-known model of Felber (66) for cytotype dynamics, we show that the transfer of genetic information between different diploid species is possible under the current understanding of interploidy mating dynamics and WGD-potentiated hybridization. A notable example of introgression, potentially explained by gene flow through polyploid bridges, pertains to genes encoding the C_4 photosynthetic pathway in *Alloteropsis semialata*, which have been laterally acquired from the polyploid *Setaria palmifolia* complex, most likely in tropical Africa, where both species co-occur (46). Most of the non- C_4 individuals in *A. semialata* are diploids, with C_4 individuals ranging from diploid to dodecaploid levels (45, 75). Although triploid hybrids have not been reported so far for *A. semialata*, the most plausible path through which genetic information coding for C_4 biochemistry could have introgressed into the diploid level of the species would be through compatible crosses that required the presence of a triploid intermediate, from which a cross of the type $3x \times 2x$ would allow the information to spread down to the diploid level. An alternative hypothesis may relate the presence of C_4 diploid individuals to the process of diploidization. However, despite the fact that diploidization is a slow process measured at macroevolutionary scales (33), there is direct experimental evidence that crosses between photosynthetic types in diploid *A. semialata* are viable (75). That is, following introgression from polyploid *S. palmifolia*, information coding for C_4 biochemistry had to necessarily travel through the different ploidy levels in *A. semialata* to reach the diploid level, where C_4 and non- C_4 individuals are then capable of exchanging genetic material.

It should be noted that the extent to which WGD facilitates interspecific introgression at the *polyploid* level is not entirely clear (21), which constrains our ability to unequivocally ascertain the impact of polyploid bridges on macro-evolutionary dynamics. For instance, in a recent study, Marburger et al. (15) were able to identify bidirectional introgression of genes controlling meiotic processes between tetraploid cytotypes of *A. arenosa* and *A. lyrata*. These shared introgressed genes, implicated in mechanisms like meiotic double strand break

formation, were deemed instrumental for the emergence of stable meiotic processes in both lineages. Nevertheless, the variability in meiotic stability found in the studied populations suggests that interspecific gene flow at the polyploid level of both species depends on optimal sorting of segregating alleles for stable lineages to emerge, which is supported by signatures of selective sweep across their genomes (15). Here, we demonstrated that although WGD-mediated interspecific introgression down to the diploid level is constrained with surging genetic divergence between mixed-ploidy populations, increased recombination rates can relieve selection pressures, thereby facilitating the influx of new allele variants into a new gene pool by increased efficiency in sorting of alleles at higher-ploidy levels. Interestingly, polyploidization has also been shown to increase recombination rates in *Arabidopsis* (76), yet a clear understanding of the relationship between genetic divergence and WGD-mediated gene flow remains elusive.

While genetic incompatibility is one of the mechanisms of reproductive isolation that is potentially circumvented by polyploidy, WGD-potentiated hybridization may also favor the transfer of genetic material across diverging populations by evading pre-zygotic barriers. Indeed, since Levin (53) postulated the principle of minority cytotype exclusion, the hypothesis that ecological disparities must exist between ploidy levels for successful establishment of polyploid cytotypes has led researchers to identify a number of cases where niche differentiation or expansion following WGD occurs (16, 37, 38, 77-79). In particular, niche expansion of polyploid organisms relative to their diploid ancestors might promote secondary contact among allopatric populations of a specific taxon, restoring therefore gene flow and counteracting reproductive isolation (21). Within this paradigm, the interpretation of the migration coefficient used in our model (κ) could span both evolutionary and ecological dimensions. On the one hand, WGD may directly breach species barriers (16, 80) by overcoming putative post-zygotic limitations, and on the other hand promote intraspecific gene flow through polyploid bridges among diverging populations, via secondary contact or niche overlaps.

For the purposes of our theoretical investigation, we neglected the possibility of triploid gamete formation by triploid cytotypes within the main manuscript. It has been shown that the frequency of triploid gametes is on average higher than the sum of diploid and haploid gametes produced by triploids (44). This will lead to an increased probability of tetraploid formation relative to the present study and would hence increase gene flow from diploid to tetraploids relative to gene flow from tetraploids to diploids (see *SI Appendix Text S6*). Despite the bias in gene flow due to differential production of gametes in triploids, the path for interploidy

bidirectional flux of genetic material should remain open. This contrasts with Cao et al.'s recent assertion (81) that crosses involving hybrid triploids invariably yields tetraploids, correlating with a reproductive barrier thwarting gene flow from tetraploids back to diploid levels, termed a "triploid block". In *SI Appendix Text S6*, we show that this effect is a consequence of the probabilities for merging different gamete ploidy levels during self-fertilization in triploid cytotypes. The model presented here is anchored in our current understanding of *interploidy* mating dynamics, and self-fertilization - though potentially important in the establishment of new cytotypes - is tangential to this study's analytical purview. The assumption that triploid cytotypes produce only diploid or haploid gametes is a simplification that allowed us to obtain an explicit analytical solution to an (extended) version of the mathematical problem posed by Felber (66). Asymmetry in the direction of gene flow due to triploid gamete formation should be considered carefully when assessing interploidy bidirectional gene flow under laboratory conditions, since it may substantially increase sampling efforts for detection of interploidy bidirectional gene flow.

Throughout this paper, we also assumed additive fitness effects that are proportional to the ploidy level. However, the hybridization of two species leading to polyploidy, or allopolyploidization, may result in nonadditive gene expression patterns on hybrids (82), which may influence fitness at the polyploid level and therefore impact the efficiency with which genetic material diffuses through ploidy levels in the recipient gene pool. The extent to which the emergence of polyploid bridges influences adaptation and disturb allele frequencies in real biological systems is yet to be determined. The flux of genes among diploid states from different species may depend on the rate at which genes flow among polyploids, which in turn may be conditioned on the spatial and temporal sequences with which higher-ploidy species come into contact. Furthermore, the fitness map used in this study imposes selection in an exogeneous manner, that is, the fitness of an organism depends on the distance of its genome to an external sequence that represents the optimum of the environment in which the individual inhabits. Selection on hybrids may as well be based on disruption of parental gene combinations, emerging therefore as the product of negative epistatic interactions. For instance, non-Mendelian segregation patterns on synthetic hybrids between polyploids from cotton have been attributed to epistatic interactions following the merging of parental genomes, whose coadapted genes constrain the genomic composition of hybrid offspring (83). Future experimental work is necessary to elucidate the effects of allopolyploidy on interspecific introgression patterns through polyploid bridges.

While polyploidy may be a transient phenomenon, i.e., an evolutionary dead-end (10), its periodic emergence in the evolutionary history of life can serve as a bridge to connect diverged taxa and redistribute genetic information among species. Here, we have provided a systematic study on the dynamics of mixed-ploidy populations and showed that, when genetic material is explicitly accounted for, WGD may promote introgression into the more (stable) diploid lineages from which polyploid species descend. Recent surveys of polyploidy and hybridization in the wild largely support the work presented in this manuscript (57). The mechanisms studied here might therefore represent an important component of reticulation during evolution and allow us to build a more general understanding of how species interactions evolve and how polyploidy affects biological networks. Further theoretical and experimental investigations on the significance of polyploid bridges may provide insightful results into the delicate interactions that weave biological networks throughout evolutionary history.

Methods

Our stochastic individual-based simulation contains a population of size N , where each discrete diploid individual is represented by a set of 2 vectors, called chromosomes, with L elements. Chromosome elements are taken from the set $\{0, 1\}$, representing biallelic loci. In the standard model, all individuals are identical, i.e., $S_{i,k} = S_{j,k}, \forall i, j \in \{N\}$ and $k \in \{L\}$. Mating is synchronous and is performed by randomly choosing two individuals, with replacement, from their corresponding mixed-ploidy population, whose gametes formed (see below) are then joined to make up the genome of the offspring, until the carrying capacity N is restored. Notice that the probability that an individual is not chosen to reproduce in each generation is $(1 - 2/N)^N$, which approaches e^{-2} in the limit $N \rightarrow \infty$, corresponding to drift. When mating is completed then bidirectional migration occurs at the tetraploid level where each individual has a probability κ of entering the next population. Each iteration, or generation, corresponds to a repetition of this process for a total of 20,000 generations.

A mating event consists of randomly selecting two individuals from the same population, with each one undergoing meiosis, where recombination among rL randomly selected loci ($r \ll 1$) takes place. In a diploid individual, meiosis pairs the two chromosomes and recombination occurs. Then, a gamete is one of the resulting chromosomes chosen with equal probability, or both chromosomes, i.e., unreduced diploid gametes, with probability v . In triploids, recombination occurs randomly between two of the three available chromosomes mimicking chromosome pairing during triploid meiosis as shown in previous experimental studies (84). In triploid meiosis three outcomes are possible: i) the individual produces no gametes with probability $(1 - \phi)$ owing to differential contribution to the gene pool due to fitness, in which case the mating event fails, ii) a haploid gamete, or iii) a diploid gamete is randomly chosen out of the three available chromosomes. We set outcomes ii) and iii) to happen with equal probability (see Eqs. 1 and 2 in Results for a mathematical description). Finally, meiosis in a tetraploid individual results in the random formation of two bivalents out of the four available chromosomes, with no preferential chromosome pairing, and a gamete is one of the bivalent pairs chosen with equal probability after recombination has occurred, following previous work (42). Although chromosome pairing in polyploid meiosis can take up different forms depending on polyploid origin (51), our goal is to isolate the effect of recombination, and not to provide a comprehensive mechanistic understanding of the possible results following differential

recombination patterns in polyploid meiosis. In *SI Appendix* Fig. S7, the reader will find an illustration detailing the meiosis procedure used in this work.

To understand the dynamics of the model under a selection regime, we developed a fitness map, assuming allele codominance, that relates the genomes of individuals to a sequence that represents the optimum phenotype (a reference genome randomly defined) in their respective mixed-ploidy populations. Let $d(G_i; G_D)$ be the total distance between the genome of individual i to its respective population's genome D (a single vector), relative to chromosome length L . Then, for each loci k in each chromosome p of individual i we define the distance to be:

$$\partial(G_{p,k,i}; G_{k,D}) = \begin{cases} \frac{1}{P}, & \text{if } G_{p,k,i} = G_{k,D} \\ 0, & \text{otherwise} \end{cases} \quad [14]$$

$$d(G_i; G_D) = 1 - \frac{1}{L} \sum_{p=1}^P \sum_{k=1}^L \partial(G_{p,k,i}; G_{k,D}) \quad [15]$$

where P is the ploidy level of individual i .

For a mating trial to be successful we require that both individuals have a fitness value larger than a random number generated with uniform distribution in the interval $(0,1)$, otherwise the system will choose other pair of individuals from the same deme until a successful mating event occurs. The fitness f of an individual is given by an exponentially decaying function of the form:

$$f = e^{-sLd(G_i; G_D)} \quad [16]$$

where s represents the selection strength that the populations shall be subject to. The genetic distance between demes is assigned upon initialization of the simulation, where the number of alleles that can differ between demes' sequences are randomly selected to match the desired genetic divergence (see *SI Appendix Algorithm S1* for implementation strategies).

Data, Materials, and Software Availability. Documentation and software to generate the results presented in this work will be made available upon publication of the manuscript. All additional data and pseudo-codes relevant for this study are available in *SI Appendix*.

ACKNOWLEDGMENTS. We thank Arthur Zwaenepoel (Unité Évolution, Écologie, Paléontologie, Université de Lille 1, Villeneuve d'Ascq, France) for helpful discussions. This work was supported by the European Research Council under the European Union's Horizon 2020 Research and Innovation program (No. 833522) and from Ghent University (Methusalem funding, BOF.MET.2021.0005.01) (to Y.V.d.P.).

References

1. M. A. Riley, M. Lizotte-Waniewski, Population genomics and the bacterial species concept. *Methods Mol. Biol.* **532**, 367-77 (2009).
2. N. H. Barton, The role of hybridization in evolution. *Mol. Ecol.* **10**, 551-568 (2001).
3. N. B. Edelman *et al.*, Genomic architecture and introgression shape a butterfly radiation. *Science* **366**, 594-599 (2019).
4. J. Steensels, B. Gallone, K. J. Verstrepen, Interspecific hybridization as a driver of fungal evolution and adaptation. *Nat. Rev. Microbiol.* **19**, 485-500 (2021).
5. D. E. Soltis, C. J. Visger, P. S. Soltis, The polyploidy revolution then...and now: Stebbins revisited. *Am. J. Bot.* **101**, 1057-1078 (2014).
6. G. P. Tiley *et al.*, Estimation of species divergence times in presence of cross-species gene flow. *Syst. Biol.* 10.1093/sysbio/syad015 (2023).
7. L. Cai *et al.*, Widespread ancient whole-genome duplications in Malpighiales coincide with Eocene global climatic upheaval. *New Phytol.* **221**, 565-576 (2019).
8. Y. Wang *et al.*, An ancient whole-genome duplication event and its contribution to flavor compounds in the tea plant (*Camellia sinensis*). *Hort. Res.* **8**, 176 (2021).
9. J. W. Clark, Genome evolution in plants and the origins of innovation. *New Phytol.* **240**, 2204-2209 (2023).
10. Y. Van de Peer, E. Mizrachi, K. Marchal, The evolutionary significance of polyploidy. *Nat. Rev. Genet.* **18**, 411-424 (2017).
11. X. Ma *et al.*, Seagrass genomes reveal ancient polyploidy and adaptations to the marine environment. *Nat. Plants* **10**, 240-255 (2024).
12. W. Bleeker, Hybridization and *Rorippa austriaca* (*Brassicaceae*) invasion in Germany. *Mol. Ecol.* **12**, 1831-1841 (2003).
13. M. H. Jorgensen, D. Ehrich, R. Schmickl, M. A. Koch, A. K. Brysting, Interspecific and interploidal gene flow in Central European *Arabidopsis* (*Brassicaceae*). *BMC Evol. Biol.* **11**, 346 (2011).
14. B. J. Arnold *et al.*, Borrowed alleles and convergence in serpentine adaptation. *Proc. Natl. Acad. Sci. U S A* **113**, 8320-8325 (2016).
15. S. Marburger *et al.*, Interspecific introgression mediates adaptation to whole genome duplication. *Nat. Commun.* **10**, 5218 (2019).
16. P. Y. Novikova *et al.*, Polyploidy breaks speciation barriers in Australian burrowing frogs *Neobatrachus*. *PLoS Genet.* **16**, e1008769 (2020).
17. D. Ståhlberg, M. Hedrén, Evolutionary history of the *Dactylorhiza maculata* polyploid complex (*Orchidaceae*). *Biol. J. Linn. Soc.* **101**, 503-525 (2010).
18. P. A. Christin *et al.*, Multiple photosynthetic transitions, polyploidy, and lateral gene transfer in the grass subtribe *Neurachninae*. *J. Exp. Bot.* **63**, 6297-6308 (2012).
19. C. Lafon-Placette *et al.*, Endosperm-based hybridization barriers explain the pattern of gene flow between *Arabidopsis lyrata* and *Arabidopsis arenosa* in Central Europe. *Proc. Natl. Acad. Sci. U S A* **114**, E1027-E1035 (2017).
20. J. Mallet, Hybrid speciation. *Nature* **446**, 279-283 (2007).
21. R. Schmickl, L. Yant, Adaptive introgression: how polyploidy reshapes gene flow landscapes. *New Phytol.* **230**, 457-461 (2021).
22. W. Albertin, P. Marullo, Polyploidy in fungi: evolution after whole-genome duplication. *Proc. Biol. Sci.* **279**, 2497-2509 (2012).
23. E. L. Breese, E. J. Lewis, G. M. Evans, Interspecies Hybrids and Polyploidy. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **292**, 487-497 (1981).
24. J. Majka *et al.*, Cytogenetic and molecular genotyping in the allotetraploid *Festuca pratensis* x *Lolium perenne* hybrids. *BMC Genomics* **20**, 367 (2019).
25. S. De Bodt, S. Maere, Y. Van de Peer, Genome duplication and the origin of angiosperms. *Trends Ecol. Evol.* **20**, 591-597 (2005).

26. D. E. Soltis, J. G. Burleigh, Surviving the K-T mass extinction: new perspectives of polyploidization in angiosperms. *Proc. Natl. Acad. Sci. U S A* **106**, 5455-5456 (2009).
27. Y. Van de Peer, T. L. Ashman, P. S. Soltis, D. E. Soltis, Polyploidy: an evolutionary and ecological force in stressful times. *Plant Cell* **33**, 11-26 (2021).
28. S. P. Otto, The Evolutionary Consequences of Polyploidy. *Cell* **131**, 452-462 (2007).
29. P. Monnahan *et al.*, Pervasive population genomic consequences of genome duplication in *Arabidopsis arenosa*. *Nat. Ecol. Evol.* **3**, 457-468 (2019).
30. G. Blanc, K. H. Wolfe, Widespread paleopolyploidy in model plant species inferred from age distributions of duplicate genes. *Plant Cell* **16**, 1667-1678 (2004).
31. A. Zwaenepoel, Y. Van de Peer, Inference of Ancient Whole-Genome Duplications and the Evolution of Gene Duplication and Loss Rates. *Mol. Biol. Evol.* **36**, 1384-1404 (2019).
32. P. Y. Novikova, N. Hohmann, Y. Van de Peer, Polyploid *Arabidopsis* species originated around recent glaciation maxima. *Curr. Opin. Plant. Biol.* **42**, 8-15 (2018).
33. A. K. Redmond, D. Casey, M. K. Gundappa, D. J. Macqueen, A. McLysaght, Independent rediploidization masks shared whole genome duplication in the sturgeon-paddlefish ancestor. *Nat. Commun.* **14**, 2879 (2023).
34. K. Vanneste, G. Baele, S. Maere, Y. Van de Peer, Analysis of 41 plant genomes supports a wave of successful genome duplications in association with the Cretaceous-Paleogene boundary. *Genome Res.* **24**, 1334-1347 (2014).
35. D. A. Levin, D. E. Soltis, Factors promoting polyploid persistence and diversification and limiting diploid speciation during the K-Pg interlude. *Curr. Opin. Plant. Biol.* **42**, 1-7 (2018).
36. Q. Bafort, T. Wu, A. Natran, O. De Clerck, Y. Van de Peer, The immediate effects of polyploidization of *Spirodela polyrhizachange* in a strain-specific way along environmental gradients. *Evol. Lett.* **7**, 37-47 (2023).
37. A. Rice *et al.*, The global biogeography of polyploid plants. *Nat. Ecol. Evol.* **3**, 265-273 (2019).
38. K. T. David, Global gradients in the distribution of animal polyploids. *Proc. Natl. Acad. Sci. U S A* **119**, e2214070119 (2022).
39. M. Ebadi *et al.*, The duplication of genomes and genetic networks and its potential for evolutionary adaptation and survival during environmental turmoil. *Proc. Natl. Acad. Sci. U S A* **120**, e2307289120 (2023).
40. I. Mayrose *et al.*, Recently formed polyploid plants diversify at lower rates. *Science* **333**, 1257 (2011).
41. D. A. Levin, Why polyploid exceptionalism is not accompanied by reduced extinction rates. *Plant Syst. Evol.* **305**, 1-11 (2019).
42. F. Kauai, F. Mortier, S. Milosavljevic, Y. Van de Peer, D. Bonte, Neutral processes underlying the macro eco-evolutionary dynamics of mixed-ploidy systems. *Proc. Biol. Sci.* **290**, 20222456 (2023).
43. B. Arnold, S. T. Kim, K. Bomblies, Single Geographic Origin of a Widespread Autotetraploid *Arabidopsis arenosa* Lineage Followed by Interploidy Admixture. *Mol. Biol. Evol.* **32**, 1382-1395 (2015).
44. J. Ramsey, D. W. Schemske, Pathways, mechanisms, and rates of polyploid formation in flowering plants. *Ann. Rev. Ecol. Syst.* **29**, 467-501 (1998).
45. J. K. Olofsson *et al.*, Low dispersal and ploidy differences in a grass maintain photosynthetic diversity despite gene flow and habitat overlap. *Mol. Ecol.* **30**, 2116-2130 (2021).
46. P.-A. Christin *et al.*, Adaptive Evolution of C4 Photosynthesis through Recurrent Lateral Gene Transfer. *Curr. Biol.* **22**, 445-449 (2012).
47. C. Phansopa, L. T. Dunning, J. D. Reid, P. A. Christin, Lateral gene transfer acts as an evolutionary shortcut to efficient c4 biochemistry. *Mol. Biol. Evol.* **37**, 3094-3104 (2020).
48. J. P. Spoelhof, R. Keeffe, S. F. McDaniel, Does reproductive assurance explain the incidence of polyploidy in plants and animals? *New Phytol.* **227**, 14-21 (2020).
49. S. P. Otto, J. Whitton, Polyploid incidence and evolution. *Ann. Rev. Genet.* **34**, 401-437 (2000).

50. T. R. Gregory, B. K. Mable, "CHAPTER 8 - Polyploidy in Animals" in *The Evolution of the Genome*, T. R. Gregory, Ed. (Academic Press, Burlington, 2005), <https://doi.org/10.1016/B978-012301463-4/50010-3>, pp. 427-517.
51. M. Cifuentes, L. Grandont, G. Moore, A. M. Chevre, E. Jenczewski, Genetic regulation of meiosis in polyploid species: new insights into an old question. *New Phytol.* **186**, 29-36 (2010).
52. J. M. Kreiner, P. Kron, B. C. Husband, Evolutionary Dynamics of Unreduced Gametes. *Trends Genet.* **33**, 583-593 (2017).
53. D. A. Levin, Minority cytotype exclusion in local plant populations. *Taxon* **24**, 35-43 (1975).
54. F. Felber, J. D. Bever, Effect of triploid fitness on the coexistence of diploids and tetraploids. *Biol. J. Linn. Soc.* **60**, 95-106 (1997).
55. B. C. Husband, The role of triploid hybrids in the evolutionary dynamics of mixed-ploidy populations. *Biol. J. Linn. Soc.* **82**, 537-546 (2004).
56. T. L. Burton, B. C. Husband, Fecundity and offspring ploidy in matings among diploid, triploid and tetraploid *Chamerion angustifolium* (Onagraceae): consequences for tetraploid establishment. *Heredity* **87**, 573-582 (2001).
57. M. R. Brown, R. J. Abbott, A. D. Twyford, The emerging importance of cross-ploidy hybridisation and introgression. *Mol. Ecol.* **33**, e17315 (2024).
58. C. C. F. Schinkel *et al.*, Pathways to polyploidy: indications of a female triploid bridge in the alpine species *Ranunculus kuepferi* (Ranunculaceae). *Plant Syst. Evol.* **303**, 1093-1108 (2017).
59. M. M. Li *et al.*, Triploid cultivars of *Cymbidium* act as a bridge in the formation of polyploid plants. *Front. Plant. Sci.* **13**, 1029915 (2022).
60. T. Peckert, J. Chrtek, Mating Interactions between Coexisting Diploid, Triploid and Tetraploid Cytotypes of *Hieracium echinoides* (Asteraceae). *Folia Geobot.* **41**, 323-334 (2006).
61. Æ. T. Thórsson, S. Pálsson, M. Lascoux, K. Anamthawat-Jónsson, Introgression and phylogeography of *Betula nana* (diploid), *B. pubescens* (tetraploid) and their triploid hybrids in Iceland inferred from cpDNA haplotype variation. *J. Biogeogr.* **37**, 2098-2110 (2010).
62. A. T. Thorsson, E. Salmela, K. Anamthawat-Jonsson, Morphological, cytogenetic, and molecular evidence for introgressive hybridization in birch. *J. Hered.* **92**, 404-408 (2001).
63. I. M. Henry *et al.*, Aneuploidy and genetic variation in the *Arabidopsis thaliana* triploid response. *Genetics* **170**, 1979-1988 (2005).
64. P. Farhat *et al.*, Gene flow between diploid and tetraploid junipers - two contrasting evolutionary pathways in two *Juniperus* populations. *BMC Evol. Biol.* **20**, 148 (2020).
65. D. Ståhlberg, Habitat differentiation, hybridization and gene flow patterns in mixed populations of diploid and autotetraploid *Dactylorhiza maculata* s.l. (Orchidaceae). *Evol. Ecol.* **23**, 295-328 (2007).
66. F. Felber, Establishment of a tetraploid cytotype in a diploid population: Effect of relative fitness of the cytotypes. *J. Evol. Biol.* **4**, 195-207 (1991).
67. Q. Qin *et al.*, Abnormal chromosome behavior during meiosis in the allotetraploid of *Carassius auratus* red var. (♀) × *Megalobrama amblycephala* (♂). *BMC Genet.* **15**, 95 (2014).
68. Y. K. Nam, D. S. Kim, Ploidy status of progeny from the crosses between tetraploid males and diploid females in mud loach (*Misgurnus mizolepis*). *Aquaculture* **236**, 575-582 (2004).
69. F. Kolar, M. Certner, J. Suda, P. Schonswetter, B. C. Husband, Mixed-Ploidy Species: Progress and Opportunities in Polyploid Research. *Trends Plant Sci.* **22**, 1041-1055 (2017).
70. A. M. Westram, S. Stankowski, P. Surendranadh, N. Barton, What is reproductive isolation? *J. Evol. Biol.* **35**, 1143-1164 (2022).
71. B. Reinhard, A survey of migration-selection models in population genetics. *Discrete Continuous Dyn. Syst. Ser. B.* **19IS - 4SP - 883**, 959 (2014).
72. R. R. Hill, Jr., Selection in autotetraploids. *Theor. Appl. Genet.* **41**, 181-186 (1970).

73. J. L. Conover, J. F. Wendel, Deleterious Mutations Accumulate Faster in Allopolyploid Than Diploid Cotton (*Gossypium*) and Unequally between Subgenomes. *Mol. Biol. Evol.* **39**, msac024 (2022).
74. D. M. Arias, L. H. Rieseberg, Gene Flow between Cultivated and Wild Sunflowers. *Theor. Appl. Genet.* **89**, 655-660 (1994).
75. M. E. Bianconi *et al.*, Contrasted histories of organelle and nuclear genomes underlying physiological diversification in a grass species. *Proc. R. Soc. Lond. B. Biol. Sci.* **287**, 20201960 (2020).
76. A. Pecinka, W. Fang, M. Rehmsmeier, A. A. Levy, O. Mittelsten Scheid, Polyploidization increases meiotic recombination frequency in *Arabidopsis*. *BMC Biol.* **9**, 24 (2011).
77. A. E. Baniaga, H. E. Marx, N. Arrigo, M. S. Barker, Polyploid plants have faster rates of multivariate niche differentiation than their diploid relatives. *Ecol. Lett.* **23**, 68-78 (2020).
78. Y. F. Molina-Henao, R. Hopkins, Autopolyploid lineage shows climatic niche expansion but not divergence in *Arabidopsis arenosa*. *Am. J. Bot.* **106**, 61-70 (2019).
79. N. Padilla-García *et al.*, The importance of considering the evolutionary history of polyploids when assessing climatic niche evolution. *J. Biogeogr.* **50**, 86-100 (2023).
80. I. Rešetnik, P. Schönswetter, M. Temunović, M. H. J. Barfuss, B. Frajman, Diploid chastity vs. polyploid promiscuity – Extensive gene flow among polyploid cytotypes blurs genetic, morphological and taxonomic boundaries among Dinaric taxa of *Knautia* (*Caprifoliaceae*). *Perspect. Plant. Ecol. Evol. Syst.* **59**, 125730 (2023).
81. Y. Cao *et al.*, Genome balance and dosage effect drive allopolyploid formation in *Brassica*. *Proc. Natl. Acad. Sci. U S A* **120**, e2217672120 (2023).
82. S. Bottani, N. R. Zabet, J. F. Wendel, R. A. Veitia, Gene Expression Dominance in Allopolyploids: Hypotheses and Models. *Trends Plant Sci.* **23**, 393-402 (2018).
83. C. X. Jiang *et al.*, Multilocus interactions restrict gene introgression in interspecific populations of polyploid *Gossypium* (cotton). *Evolution* **54**, 798-814 (2000).
84. M. L. Becak, W. Becak, Further studies on polyploid amphibians (*Ceratophrydidae*). 3. Meiotic aspects of the interspecific triploid hybrid: *Odontophrynus cultripes* (2n=22) x *O. americanus* (4n=44). *Chromosoma* **31**, 377-385 (1970).

FIGURES

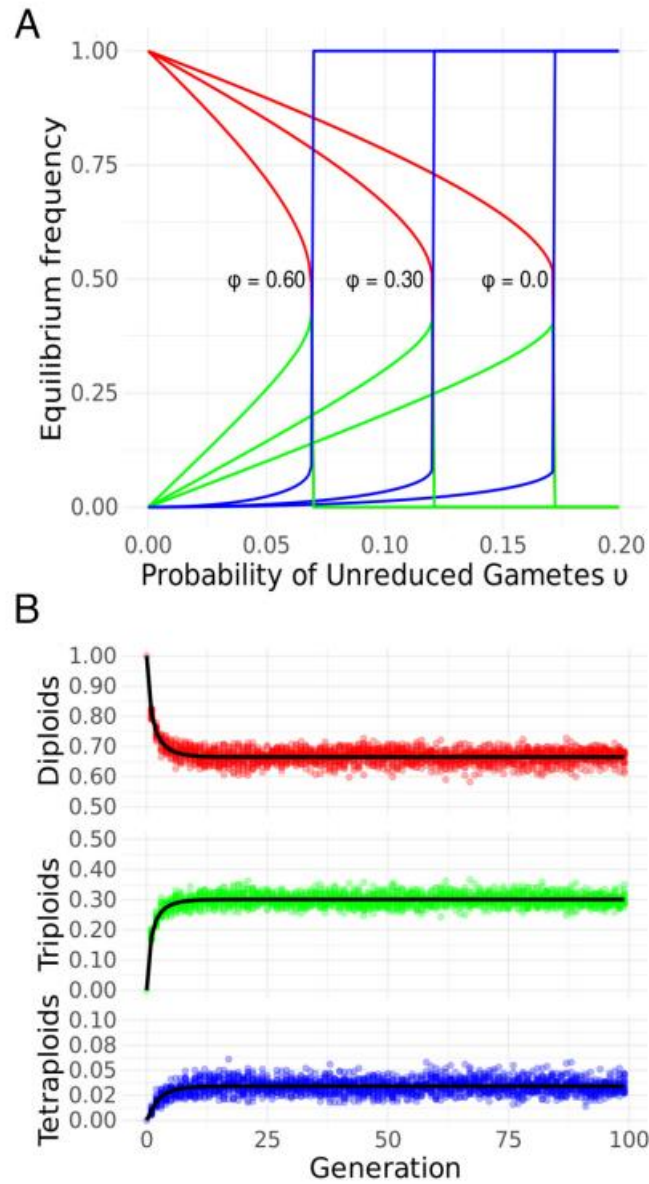


Fig. 1. Dynamics of cytotype frequencies in a single deme. (A) Equilibrium frequency of diploids (solid red line), triploids (solid green line) and tetraploids (solid blue line) as a function of the probability of unreduced gametes in diploids v , for triploid fitness $\varphi = 0$, $\varphi = 0.30$ and $\varphi = 0.60$. Maximum frequency of unreduced gametes \bar{v} below which coexistence is verified corresponds to the vertical lines, where tetraploids overtake the system. (B) Frequency of each cytotype in the mixed-ploidy population per generation, with $\varphi = 0.30$ and $v = 0.10$, given by the set of Eqs. 1 and 2 in solid black lines, and by the individual-based stochastic version of the model (in faintly colored filled circles) with finite population size $N = 1000$ (see *Methods*).

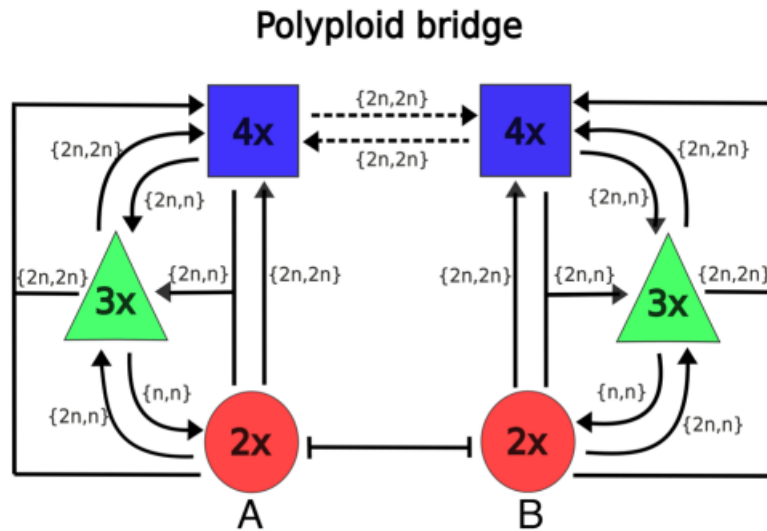


Fig. 2. Mixed-ploidy model of interspecific gene flow. The intercytotype gene flow within and between mixed-ploidy populations A and B is depicted, with diploids (2x red circle), triploids (3x green triangle) and tetraploid (4x blue square) cytotypes, following largely independent dynamics, except by migration at the tetraploid level, represented by the dashed lines connecting the tetraploid cytotypes at the upper part of the picture. Solid lines connecting cytotypes represent possible intercytotype mating events, with the direction of the arrow indicating the resulting ploidy level. The unordered set next to each line represents the gametes involved in the mating trial. For example, a 2n (diploid) gamete flowing from a 4x cytype and a n (haploid) gamete flowing from a 2x cytype, *i.e.*, {2n,n}, will result in a 3x cytype. The solid line between the 2x cytypes from populations A and B represents the species barrier.

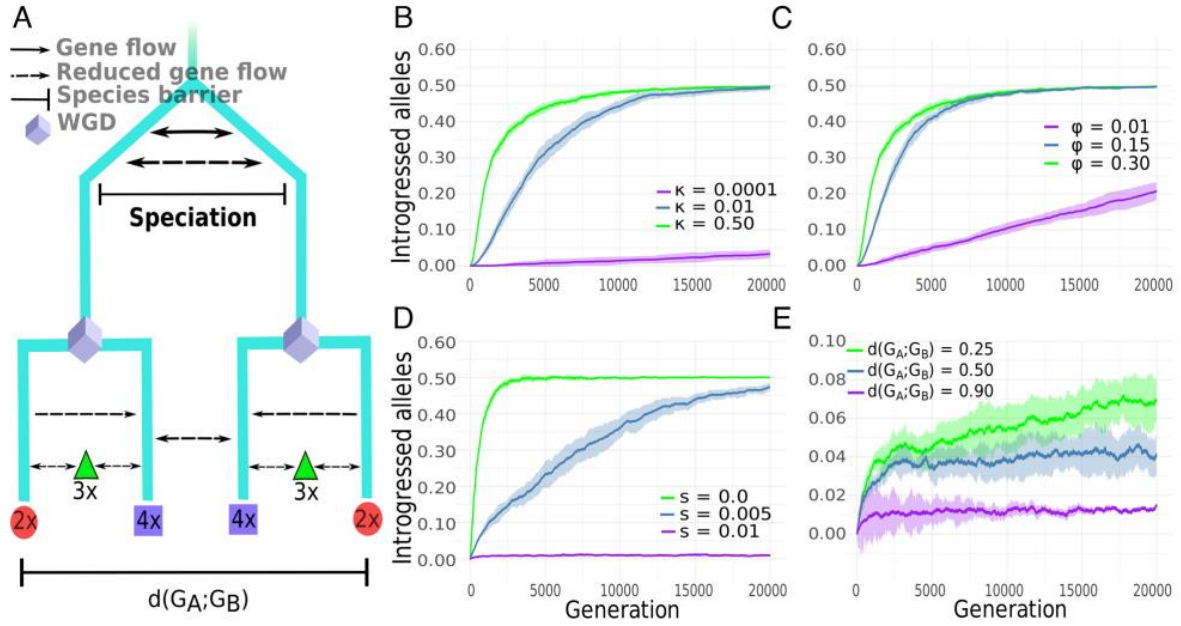


Fig. 3. Frequency of introgressed alleles at the diploid level per generation. (A) The evolutionary history of a hypothetical lineage splitting into two different species, A and B, is presented with each one displaying WGD events (following (21)). Diploids (red circle), triploids (green triangle) and tetraploids (blue square), with the possible gene flow pathways as studied are highlighted. The genetic distance, $d(G_A; G_B)$, represents the nucleotide polymorphisms between the two species following their evolutionary history (see *Methods*). (B and C) The average frequency of homozygous loci with introgressed alleles per generation is shown, for different values of κ and ϕ , respectively, and $\phi = 0.30$ (B) and $\kappa = 0.50$ (C). Homozygous and heterozygous loci are distinguished to highlight that fixation proceeds iteratively through successive recombination events, until the system converges to a 50% equilibrium point of introduced and native homozygous loci in each deme (see *SI Appendix Fig. S5* for the evolution of introgressed alleles considering only heterozygous loci). (D and E) The total frequency of introgressed alleles per generation for different selection strength values s (with $d(G_A; G_B) = 100\%$) (D) for alternative genetic distances using a selection strength of $s = 0.01$ (E). Shaded ribbons represent ± 1 standard deviation from the average.

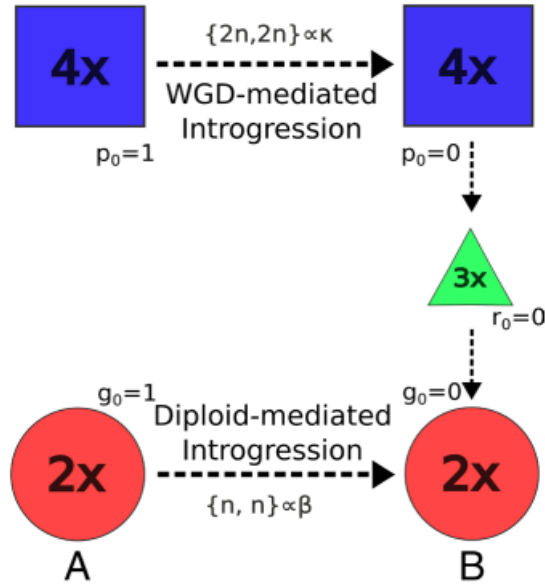


Fig. 4. Routes for adaptive introgression in a new diploid gene pool. Two routes are presented for unidirectional gene flows from species A to species B. Species A is initially ($t = 0$) fixed for a beneficial allele whose frequency in the tetraploid ($4x$) and diploid ($2x$) states is given by p_t and g_t , respectively. Gene flow from species A to species B is represented by the transfer of gametes (in curly brackets) between gene pools as before. Gene flow at the diploid level occurs with rate β and at the tetraploid level with rate κ . The speed with which a beneficial allele will introgress in the diploid gene pool of species B (lower left corner) was compared when genetic information flows interspecifically through the route $4x \rightarrow 4x \rightarrow 3x \rightarrow 2x$, with the case where gene flow occurs through the route $2x \rightarrow 2x$, mediated by β .

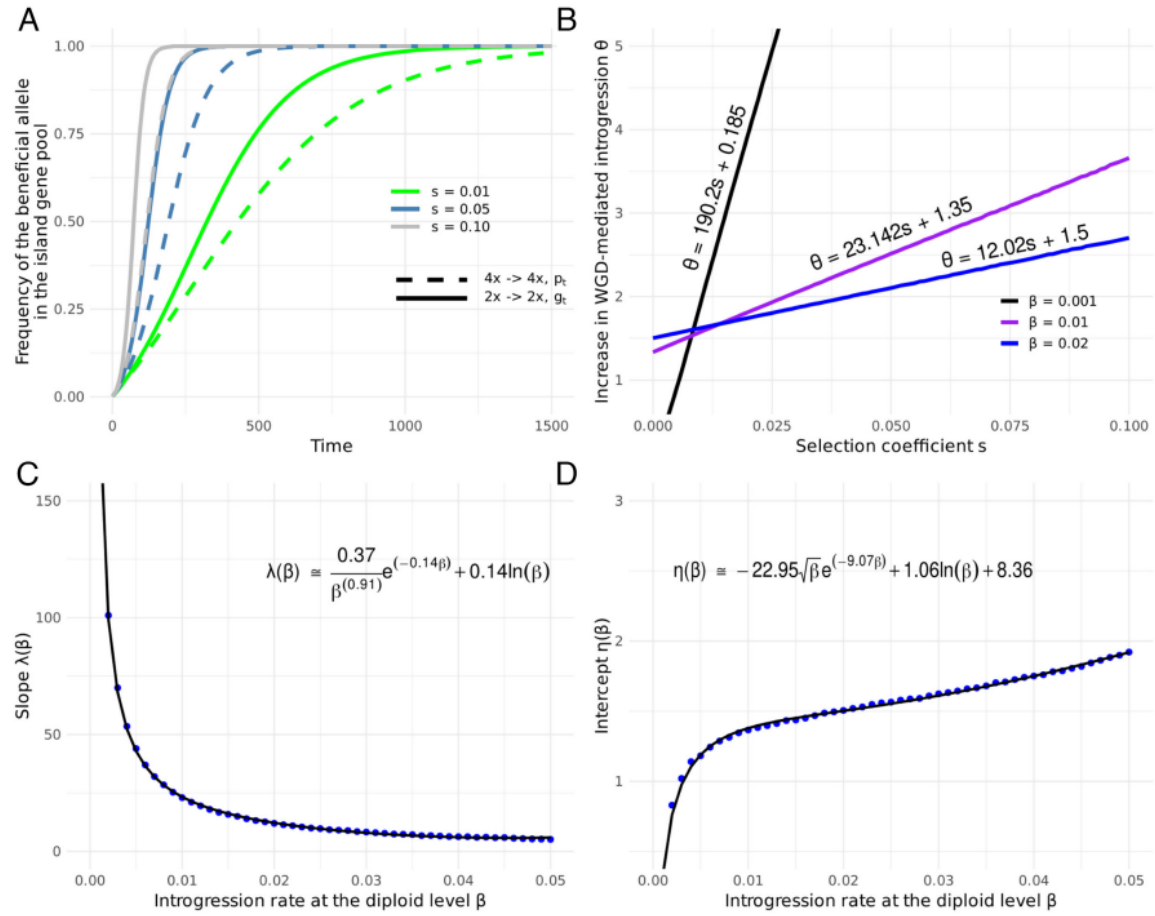


Fig. 5. Adaptive introgression efficiency comparison between gene flow mediated by diploids and tetraploids in the mainland-island model. (A) Frequency of the beneficial allele A as a function of time (iterations) and selection strength s in the island at the diploid level through unidirectional gene flow from the mainland's diploid population (g_t , solid lines), and at the tetraploid level through unidirectional gene flow from the mainland's tetraploid population (p_t , dashed lines). Gene flow rates are the same in both cases: $\beta = \alpha = 10^{-3}$. Additionally, $\varphi = 0$. (B) Minimum increase in WGD-mediated introgression rate θ to achieve fixation of the beneficial allele within the same time interval as achieved by diploid-mediated introgression ($\varphi = 0.30$). (C and D) Relationship between the slopes and intercepts of the linear equations (B) as a function of introgression at the diploid level β . Equation coefficients were retrieved by the Levenberg-Marquadt algorithm.