Interactions between mechanisms of reproductive isolation

Alexandre Blanckaert[a*] and Vitor C. Sousa[a]

[a] cE3c - Centre for Ecology, Evolution and Environmental Changes & CHANGE - Global Change and Sustainability Institute, Faculdade de Ciências, Universidade de Lisboa, Campo Grande, Portugal

Corresponding author *: blanckaert.a@gmail.com

Conflict of interest: The authors declare no conflict of interest.

Funding: This work was funded by the strategic project UIDB/00329/2020 granted to CE3C from the Portuguese National Science Foundation Fundação para a Ciência e a Tecnologia (FCT). This study was funded by HFSP Young Investigator Grant RGY0081/2020 to V.S., FCT Grant 2022.03475.PTDC to A.B. and V.S. and FCT Grant DOFLOW to V.S.. We thank the National Network for Advanced Computing (RNCA) and INCD (incd.pt) for access and use of their computing infrastructure, funded by FCT to A.B. and V.S. (2022.38032.CPCA.A0 and 2022.15858.CPCA.A2).

Author contributions: A.B. and V.S. designed the study, A.B. wrote the simulation code and conducted the analysis, and A.B. and V.S. wrote the manuscript.

Acknowledgments: We thank João Carvalho, Inês Fragata and the Evolutionary Genomics and Bioinformatics lab for discussion of the project and Claudia Bank for comments on the manuscript.

Abstract

Speciation is responsible for the diversity of species observed today and corresponds to the build-up of reproductive isolation between populations. Reproductive isolation can be generated by different mechanisms that have been extensively characterized, yet how their interactions affect speciation remains largely unknown. Here, we explicitly model the interaction of three key mechanisms (local adaptation, mate choice and genetic hybrid incompatibilities) quantifying their relative contribution to the evolution of reproductive isolation. We modeled two populations exchanging migrants using Fisher Geometric Model for local adaptation, phenotype matching for mate choice, and multiple pairs of Bateson-Dobzhansky-Muller Incompatibilities (DMI). All three mechanisms were determined by the same set of loci, creating conditions for interactions between barriers both at the genetic and population levels. We found no cases where the three barriers evolved. Instead, two barriers could evolve depending on the migration rate: either local adaptation and genetic incompatibilities for limited migration, or local adaptation and mate choice for higher migration. Our results showed that local adaptation due to ecological differentiation was the first to evolve and by far the most effective reproductive barrier. Finally, we demonstrated that in a polygenic model, populations could become locally adapted and evolve strict mate choice, yet they would not accumulate incompatibilities provided that there was sufficient gene flow.

Speciation has been a central theme in evolutionary biology since its early origin Darwin [1859]. Yet, despite a truly large body of work, general rules of how speciation works and which mechanisms are involved remain to a large extent a mystery. So far, most findings provide a case by case set of rules rather than universal ones (see Table 1 in Sobel et al. [2010]). Two key questions have been under active research and debate. First, what should be measured to capture progress in the speciation process (assuming such process is linear) Westram et al. [2022], Mallet and Mullen [2022]? Second, when is speciation complete/what is the end of the speciation process? This was historically defined as complete reproductive isolation between the two diverging populations Coyne and Orr [2004]. However, recent advances in genomics reveal that gene flow between well-defined species is far from uncommon Payseur and Rieseberg [2016], making it difficult to draw the line of when speciation completes Hey and Pinho [2012], Dopman et al. [2024]. Furthermore, it has been recently argued that partial reproductive isolation may not be a transient state on the way to speciation but that it may also be "a final equilibrium" state Servedio and Hermisson [2020], Barraclough [2024], further blurring the notion of species.

Different mechanisms of reproductive isolation can drive the speciation process and they have been classified depending on the timing (pre- or postzygotic) and nature of selection acting upon the mechanism of reproductive isolation, whether it is due to constrains at the genomic level (intrinsic) or imposed by the environment (extrinsic) Coyne and Orr [2004]. We can therefore interrogate and disentangle which reproductive isolation process(es) have the largest impact during speciation and whether their interaction is synergistic (leading to coupling of genetic barriers - sensu Butlin and Smadja [2018]) or antagonistic. There have been many attempts to characterize the joint evolution of different mechanisms of reproductive isolation (*i.e.*, mate choice, hybrid load), or the evolution of one conditioned on the other, with assortative mating models via "magic trait" or reinforcement being the more widely known examples. For example, Maisonneuve et al. [2024] showed that, in the case of unviable F1s, mate choice will only evolve towards targeting a neutral discriminant trait over a trait providing a fitness advantage when the risk of choosing the wrong mate is too high, indicating that adaptation and mate choice may not act synergistically.

There has been a recent push to fully characterize the different mechanisms (or a large proportion of) acting between pairs of nascent species Cruz et al. [2024], Aguillon et al. [2024]. However, quantifying the role of these mechanisms in natural populations can be challenging, even for species amenable to laboratory manipulation (e.g., Aguillon et al. [2024]). In addition, there are no general rules emerging. For example, Castillo et al. [2015] conducted experimental evolution of *Caenorhabditis remanei* on different substrates, measuring the evolution of local adaptation, mate choice and genetic incompatibilities, concluding that local adaptation was not relevant for the evolution of reproductive isolation. This contrasts with other studies highlighting the importance of local adaptation in natural populations (reviewed in Rundle and Nosil [2005], McGee et al. [2020]). Moreover, it was argued that partial reproductive isolation may represent an optimal strategy Servedio and Hermisson [2020], Barraclough [2024]. Indeed, the coupling of

RI barriers, an important step in the speciation process Barton [1983], Butlin and Smadja [2018], Dopman et al. [2024], may lead to overall weaker rather than stronger reproductive isolation Aubier et al. [2024].

Historically, speciation has been studied depending on its spatial context, focusing on the dichotomy between sympatry and allopatry, but there is a recent shift towards considering parapatry Gavrilets [2003, 2014], which encompasses the whole continuum, with allopatry and sympatry as extreme cases. Yet, several authors proposed to consider whether speciation occurs with or without gene flow, irrespective of the spatial context Butlin et al. [2008], Stankowski and Ravinet [2021], Bolnick et al. [2023]. When there is gene flow due to migration between populations, genetic incompatibilities create a hybrid load that may act as a sieve, leading to removal of genetic incompatibilities from both populations. Thus, provided that the mutations involved in incompatibilities do not confer local selective advantage, or are not tightly linked to a mutation that does, we expect genetic incompatibilities to be lost when there is gene flow Bank et al. [2012], Blanckaert and Hermisson [2018]. In addition, migration between two populations in contact with each other creates divergent selection on mate choice traits McPeek and Gavrilets [2006], Yamaguchi and Iwasa [2013] to avoid the production of unfit hybrids (whether the cause is intrinsic or extrinsic).

Here, we characterized how local adaptation, mate choice and genetic incompatibilities interact with each other, quantifying their impact on the strength of reproductive isolation between two populations evolving with migration. We investigated whether the evolution of one mechanism impacts the evolution of the others, and quantified their interaction by assessing their joint effect on reproductive isolation, compared to their individual contributions. We focus on three specific mechanisms of reproductive isolation, each standing for a different type according to the classification mentioned above: mate choice (here an intrinsic prezygotic mechanism), accumulation of genetic incompatibilities (intrinsic postzygotic) and local adaptation (extrinsic pre- and postzygotic). In particular, we designed a model where reproductive mechanisms interact with each other at the genetic level: a large number of loci contribute to the phenotype of the individual, with a subset forming intrinsic genetic incompatibilities. Moreover, the fitness of an individual is defined as the compound effect of phenotype (i.e., how well adapted it is to the environment) and of intrinsic genetic load due to genetic incompatibilities. Finally, the phenotype of each individual also affects its attractiveness to potential mates in the population (mate choice by phenotype matching). Under this model, we showed that local adaptation is the first mechanism to evolve and also the key mechanism responsible for the strength of reproductive isolation. Mate choice and the accumulation of genetic incompatibilities seem to be mostly exclusive mechanisms, and which one acts complementary to local adaptation is mediated by the level of migration between populations. Interestingly, we find that even populations that evolve local adaptation and strict mate choice, might not accumulate genetic incompatibilities, provided that there is sufficient gene flow.

Model

Population model

We used an island-island model and considered the long term evolutionary ($> 10N$ generations) fate of 2 populations of N diploid dioecious individuals exchanging migrants at rate m. The two islands are referred to as environments A and B, and for simplicity, we refer to the population residing in each environment as population A and B, respectively.

To explicitly consider the interaction of the three barriers types, we considered a genome where loci determining the phenotype, which affects local adaptation and mate choice, could also be involved in genetic incompatibilities, as illustrated in Figure 1. The overall fitness of an individual resulted from the combined effects of intrinsic and extrinsic fitness components. The extrinsic fitness depended on the phenotype of a given individual and its environment. Importantly, a subset of the loci that code the phenotype (i.e., involved in local adaptation) also formed pairs of hybrid incompatibilities, determining the intrinsic fitness. Mate choice was modeled as assortative mating (in its stricter sense, i.e., positive assortative mating; Jiang et al. [2013]), with individuals preferring mates with phenotypes similar to their own, depending on genotypes at two modifier loci.

The life cycle of both populations at each generation included the following sequence of events: (i) selection, (ii) migration, (iii) mating and (iv) zygote formation (including recombination and mutation). We considered two possible life cycles differing in whether selection or migration occurred first. In the "adult dispersal" selection occurred before migration, whereas in the "juvenile dispersal" selection occurred after migration. Since fitness of an individual was computed when individual reached adulthood, the difference between the two life cycle boiled down to whether immigrating individuals had their fitness determined by the environment they grew up in ("adult dispersal") or by the one they inhabited at the time of reproduction ("juvenile dispersal"). Migration was modeled as a Poisson process, of parameter $N \times m$. Population size N was assumed to be the same in both populations and constant when forming the next generation. Each population, regardless of the number of adults, produced N offspring with a 50:50 sex ratio.

Genotype and phenotype: The genome of each diploid individual was composed of three different (autosomal) elements: one large genomic element determining the phenotype and genetic incompatibilities, composed of n_{LB} linkage blocks with $\frac{L}{n_{LB}}$ biallelic loci each, and two modifier loci determining mate choice. We assumed that the phenotype for n_t quantitative traits was determined by the genotype at the L biallelic loci. For the L biallelic loci, mutation was bidirectional, symmetric and happened at rate μ (per individual, per loci, and per generation). Recombination between adjacent loci (within a block) happened at rate ρ (per individual and per generation), while recombination between linkage blocks happened at rate ρ_{LB}. In the special case of $\rho_{LB} = 0.5$, the large genomic element was effectively split into n_{LB} independent chromosomes. All mutations were pleiotropic and isotropic (affecting all n_t phenotypic traits equally), with phenotypic effects $(z_{i1}, .., z_{in_t})$ for the n_t traits drawn from a multivariate centered normal distribution with the covariance matrix $\Sigma = \sigma_\mu^2 I_{n_t}$, where I_{n_t} is the identity matrix and σ_μ^2 the variance

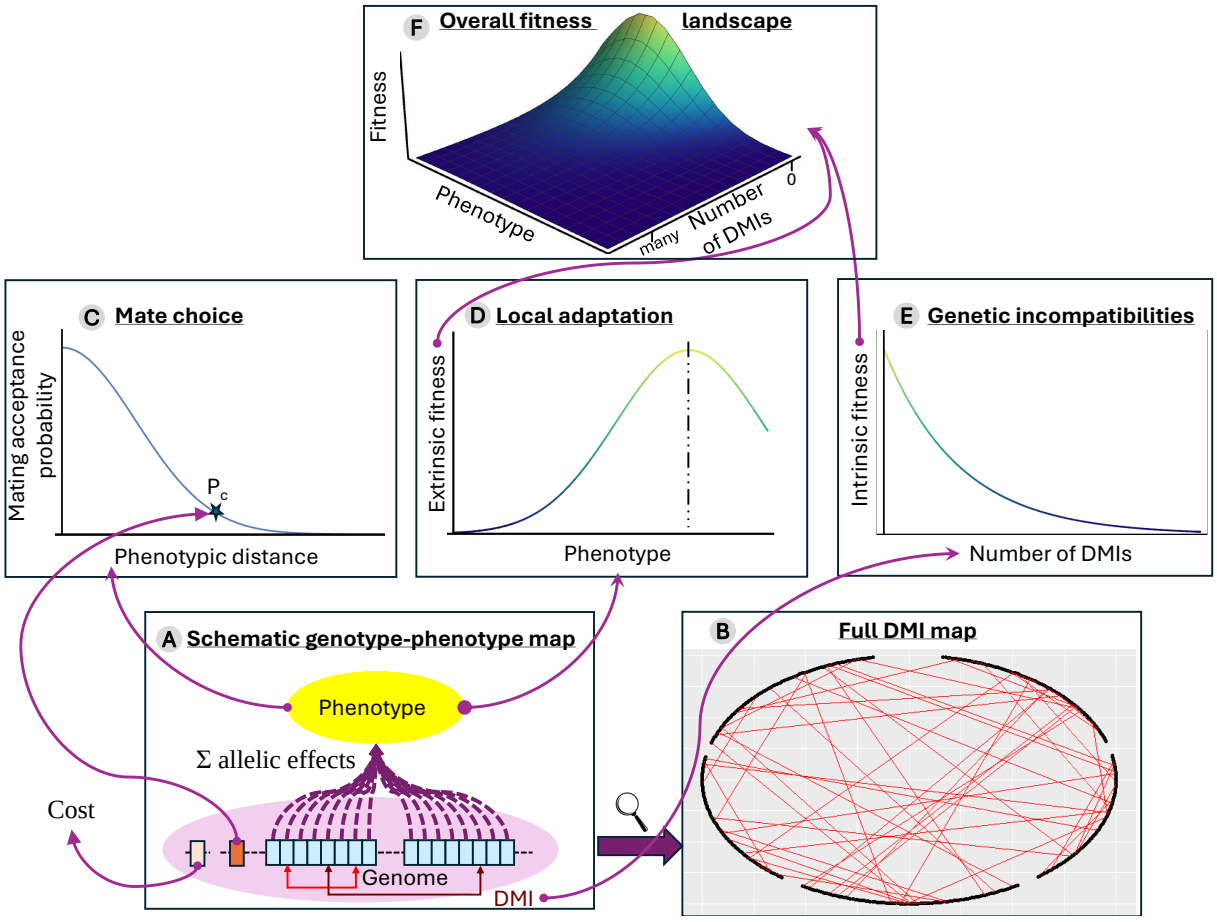

Figure 1: Schematic illustration of the genetic basis of the different mechanisms of reproductive isolation (RI). The diploid genome (panel A) consists of L loci separated in n_{LB} linkage blocks depicted by the blue squares. Their effect on each phenotypic trait are summed to generated the phenotype, depicted in yellow. In addition, a modifier locus (in dark orange) controls the strength of preference (P_C), which determines how steep the probability of accepting a mating function is (panel C, see eq. (4)). Finally, a second modifier locus (in light orange) controls the cost of preference (χ). The phenotype determines the extrinsic fitness through a quadratic function (panel D, see eq. (1)). Finally, genetic incompatibilities are depicted by the red arrows between blue squares (panel B for the full map for the default case), and their number determined the intrinsic fitness (panel E, see eq. (3)). The overall fitness is obtained by combining extrinsic and intrinsic fitnesses, as depicted in panel F.

of mutational effects. Mutational effects were additive at the phenotypic level, hence the phenotype of a (diploid) individual was given by $Z_j = \sum_{i=1}^{L} X_i z_{ij}$, for a given phenotypic trait j, with X_i taking the value 0 if the individual was homozygous for ancestral allele, 1 if heterozygous and 2 if homozygous for the derived allele at locus i. This implies that we assume all loci are additive at the phenotypic level. For simplicity, we will refer to the phenotype that is at the optimum in the focal environment as the "optimal phenotype", and to the phenotype at the optimum in the other environment as the "alternative phenotype".

We assumed that preference (P_C) varies between 0 (strict assortative mating) and 1 (no preference), and that the cost of preference (χ) varies between 1 (strong cost of preference) and $+\infty$ (no cost). For simplicity, we considered a single preference locus acting as a modifier locus [Haldane, 1941, Nei, 1967, Feldman et al., 1996] with multiple alleles, allowing for trait values to vary between 0 and 1 (or > 0 for the cost locus), rather than a quantitative trait with many underlying loci. We also modeled the cost of preference with a modifier locus, for identical reasons. Mutation happened at rate 10μ for the two modifier loci (preference and cost) and the new allele effects were obtained as deviations from the trait value of the current allele. Deviations were drawn from a normal centered distribution with respective standard deviation σ_p and σ_c. The effect of alleles was additive, and therefore for preference the trait value an allele could take was bounded between 0 and 0.5, and for the cost locus, the allelic trait value had a minimum of 0.5. For both preference and cost loci, allele values exceeding the boundaries upon mutations were set to the corresponding boundary value. The preference locus was situated at the beginning of the main genomic element, at a genetic distance ρ_{MC} from the first position of the first linkage block, while the cost locus segregated independently from all other genomic elements.

Given that the range of values that the phenotype and mate choice could take was bounded, these traits suffered from a mutational bias: when the phenotypic trait (or the mate choice allele effect) reached values close to their extrema, the mean displacement caused by new mutations was, on average, no longer 0 but biased towards the other extremum (Fig. S1).

Finally, sex was randomly assigned to each individual at birth according to the expected sex ratio (50:50).

Fitness: The fitness w of an individual was obtained by considering both extrinsic w_{ext}^{A} (due to the environment, here A) and intrinsic (due to the genetic incompatibilities) w_{int} components: $w = w_{ext}^{A} * w_{int}$, and corresponded to the relative probability of an individual to initiate a mating. Note that initiating a mating did not imply reproduction, since preference and cost of preference may impact the outcome.

Local adaptation (extrinsic fitness): The extrinsic component of fitness was modeled using Fisher Geometric Model Fisher [1930]. The (absolute) fitness of each individual depended on the (euclidean) distance between their phenotype, given by the $(Z_1, .., Z_{n_t})$ vector over the n_t phenotypic traits, and the environmental optimum. The relation between phenotype and fitness was obtained from Gros et al. [2009] with a slightly

different choice of parametrization, to explicitly account for fitness in the alternative environment. The fitness of an individual in environment A is given by:

$$w_{ext}^{A} = e^{log(\Omega^B)\left(\frac{\Sigma_{j=1}^{n_t}(Z_j - Z_j^A)^2}{\Delta_{AB}}\right)^{\frac{Q}{2}}} \tag{1}$$

where Z_j is the phenotypic value of trait j, Z_j^A is the optimal phenotypic value for trait j in environment A, Δ_{AB} is a constant corresponding to the squared (Euclidean) distance between the phenotype of two individuals at the optima in environment A and environment B ($\Delta_{AB} = \Sigma_{j=1}^{n_t}(Z_j^A - Z_j^B)^2$), Q ($Q \in [0, \infty[$) controlled the average epistasis between two mutations at the optimum, and Ω^B is a constant corresponding to the fitness of an individual in environment A with a phenotype perfectly adapted to environment B (i.e., $Z_j = Z_j^B$). In the fitness landscape Gros et al. [2009] when $Q < 2$, epistasis was positive on average, otherwise it was negative (with null epistasis for $Q = 2$). The fitness in environment B (w_{ext}^B) was obtained by substituting Z_j^A by Z_j^B, and Ω^B by Ω^A. Based on equation (1), we could determine when F_1 individuals, from parents at the phenotypic optimum at the two environments (e.g., mother with phenotype Z_j^A and father Z_j^B), displayed hybrid vigor (their fitness was greater than the mean of the parents' fitness; given by $\frac{1+\Omega^A}{2}$). Assuming that F_1 individuals had an intermediate phenotype (i.e., $\forall j, Z_j = \frac{Z_j^A + Z_j^B}{2}$, leading to $w_{ext}^{A,F_1} = e^{0.25^{Q/2} log(\Omega^B)} = (\Omega^B)^{2^{-Q}}$; by symmetry $w_{ext}^{B,F_1} = (\Omega^A)^{2^{-Q}}$), whether hybrid vigor was observed mainly depended on the mean epistasis Q of the landscape (and the parents' fitness):

$$w_{ext}^{A,F_1} > \frac{1+\Omega^B}{2} \text{ if } Q > -\frac{\log\left(\frac{\log\left(\frac{\Omega^B+1}{2}\right)}{\log(\Omega^B)}\right)}{\log(2)} \tag{2}$$

Under the default parameters (Table 1), weak hybrid vigor was therefore the default behavior.

Genetic incompatibilities (Intrinsic fitness): We modeled intrinsic fitness, as pairs of Bateson-Dobzhansky-Muller incompatibilities (DMIs) Bateson [1909], Dobzhansky [1936], Muller [1942]. Importantly, the loci involved in incompatibilities also affected the phenotype and therefore could be involved in local adaptation, meaning that the two mechanisms may have the same genetic loci involved. In details, we considered n_{DMI} pairs of incompatibilities, with the loci chosen among the L loci that affected the individual phenotype. Per default, each locus could form at most a single DMI (sampling without replacement), with the two interacting loci chosen independently of their position along the genome. Each DMI was codominant and their interaction multiplicative, resulting in the following formula for the intrinsic fitness component:

$$w_{int} = \Pi_{k=1}^{n_{\text{DMI}}}(1 - \epsilon^{X_k * Y_k}) \tag{3}$$

where X_k and Y_j are the counts of the number of derived alleles (0, 1 or 2) at the two loci forming the DMI. As an extension, we considered a more interconnected network,

where a single locus may form multiple DMIs, however those incompatibilities remain pairwise, and no higher order epistasis was considered. Finally, we also investigated a scenario, named "optimized" (only for $n_t = 1$) where exactly half of loci ($L/2$) affected the phenotypic trait (Z_1) positively (favored in environment A) and the other half ($L/2$) affected the phenotype negatively (favored in environment B). Mutations that affected the phenotype positively, formed pairwise DMIs with mutations affecting the phenotype negatively, with each mutation only involved in a single DMI. This scenario, while unrealistic, represented an interesting limiting case with regards to the accumulation of DMIs, as it minimizes the possible genetic conflict between ecological divergence and genetic incompatibility.

Mate choice: We modeled mate choice using a phenotype matching approach following Irwin [2020], where the probability of accepting a mate depended on the phenotypic distance between the two individuals, and was given by a Gaussian function. Rather than defining a variance, the shape of the Gaussian function was controlled by one parameter, P_c, that corresponded to the probability of an individual with the "optimal phenotype" accepting a mate with the "alternative phenotype". For simplicity, we assumed that males were the choosy individuals, but our results hold for female mate choice. Mate choice proceeded as followed: first a random pair of individuals was chosen, according to their respective fitness. Second, the male accepted the potential female depending on his preference (P_c) and how phenotypically similar the two potential mates were, with probability

$$P(mating) = \begin{cases} 1 & \text{if } P_c = 0 \text{ and } \Sigma_{j=1}^{n_t}(Z_j^m - Z_j^f)^2 = 0 \\ e^{log(P_c)\frac{\Sigma_{j=1}^{n_t}(Z_j^m - Z_j^f)^2}{\Delta_{AB}}} & \text{else} \end{cases} \tag{4}$$

where Z_j^m and Z_j^f are the phenotypes at trait j of males and females, respectively. If the potential mate was rejected, the male was kept and a new potential female was chosen according to its fitness. If a male failed to accept a female in χ attempts, the male was discarded (but available to be chosen again in the first step). Mating resulted in the production of a single offspring. This process continued until N offspring were produced.

Starting conditions and population metrics: We considered three possible starting conditions:

- "Ancestral" state: both populations started identical and were fixed for allele 0 at all loci.
- Secondary contact state, "SC": the two populations started differently fixed for each DMI locus, and further mutations were included such that $\forall j$, with j being a given phenotypic trait, $Z_j \approx Z_i^A$ for environment A (and $Z_j \approx Z_i^B$ for environment B).
- Secondary contact with initial mate choice state, "SC + initial mate choice": the starting conditions were identical to the "SC" state but with preexisting strict mate choice ($P_C = 0.001$).

We kept track of the population by computing the mean absolute fitness, as well as its extrinsic and intrinsic components, the mean phenotype, the mean preference and the hybrid load between populations (defined below). All metrics were computed for adults individuals after selection and migration had already occurred. In particular, this meant that if all individuals of the two populations were at the phenotypic optimum at A (Z_j^A) and B (Z_j^B), the mean phenotype for trait j of population A was given:

$$\bar{Z}_j = (1-m)Z_j^A + mZ_j^B \tag{5}$$

The hybrid load was defined as the reduction in (intrinsic) fitness caused by genetic incompatibilities, quantified by computing the intrinsic fitness component of simulated ephemeral (not part of the next generation) F_1 individuals resulting from the cross of individuals from the two environments, correcting by the mean intrinsic fitness to their parents ($\bar{w}_{int}^A$ and $\bar{w}_{int}^B$). This correction allows us to disentangle the hybrid load (defined as the reduction in fitness due to genetic incompatibilities present in different populations) from the mutational load caused by DMIs within each population. The ephemeral F_1 individuals were obtained with males of population A (resp. B) choosing females among population B (resp. A), according to their preference and preference cost. Given that we were looking at F_1 individuals, all DMIs between populations would be in the double heterozygote state. Therefore, we computed a number of equivalent DMIs needed to generate the same hybrid load (corrected by within population mutational load due to genetic incompatibilities), in the absence of polymorphism:

$$n_{\text{DMI}}^{eq} = \frac{log(\frac{2\bar{w}_{F_1}^{int}}{\bar{w}_{int}^A + \bar{w}_{int}^B})}{log(1-\epsilon)} \tag{6}$$

where $\bar{w}_{F_1}^{int}$ was the mean intrinsic fitness of hybrids, capturing how many pairs of DMIs isolated the two populations from each other. In scenarios with migration, this number of DMIs could be underestimated since populations would contain a proportion of individuals with the alternative phenotype. Hence, immigrant males would prefer to mate with those despite their lower extrinsic fitness, reducing the generation of $F_1 s$ from individuals from the two environments. This effect could be further exacerbated by the life cycle, when selection preceded migration.

Unless specified otherwise, we used a single genetic background (i.e., map of genetic interaction and mutational effects) for a given set of (relevant) parameters (i.e., for $\{L, n_t, n_D\}$). In particular, we used the following subsets:

- "default" map: 50 DMIs pairs among 500 loci, affecting the phenotype, which was defined as a single phenotypic trait. Each locus could be involved in at most a single DMI pair. This is the map illustrated in Figure 1B.

- "default 3D" map: 50 DMIs pairs among 500 loci, affecting the phenotype, which was defined as three phenotypic traits. Each locus could be involved in at most a single DMI pair.

- "optimized" map: 250 DMIs pairs among 500 loci. The 250 loci affecting the phenotype positively interacted in a pairwise and unique way with the 250 loci affecting the phenotype negatively. The phenotype consisted of a single trait.

- "network 50L" map: The genome was formed by 800 loci; half of them , chosen randomly affected the phenotype. 50 loci were randomly picked among the other 400 (neutral) loci and formed 200 DMIs pairs. The phenotype consisted of a single trait.

Measuring Reproductive Isolation (RI): reduction in gene flow caused by each RI mechanism

To measure the reduction in gene flow between the two populations, we considered the fate of two neutral markers, one freely recombining and the other situated in the middle of the genome. One individual (refereed to as focal) from the population A was transferred to population B. We follow the fate of these neutral markers (introduced in 2 copies) over a maximum of N generations. The fate of these alleles was always tracked with the 2 populations in isolation, regardless of the initial migration rate. We chose this design to avoid the neutral marker surviving in the alternative population through back migration. Although invasion probability was the most obvious metric to track, it was computationally heavy, especially when considering large populations. Therefore, we measured the time of loss of the neutral markers or their frequency after N generations, if they were still polymorphic. By default, we considered the fate of 1,000 invasion events (each time choosing the focal individual randomly) in populations of $N = 10,000$ individuals. We considered all the three Pythagorean means: arithmetic, geometric and harmonic, as the first one is known to be sensitive to rare large values.

To assess the impact of each mechanism of RI, we measured the strength of RI under all possible combinations that could have influenced their evolution. For the general case with the three barriers: local adaptation, mate choice and accumulation of genetic incompatibilities, there were eight different possible combinations of mechanisms of RI (referred to as "RI barriers" for the rest of manuscript). If a population did not evolve in presence of a given mechanism (for example mate choice), then we do not consider this mechanism when measuring RI. Therefore, we only have four possible "RI barriers" if two mechanisms were active (none, RI1, RI2, RI1+RI2) and only two "RI barriers" if only one mechanism was active (none, RI1).

To evaluate how the reduction of gene flow depended on the focal individual, we measured its fitness components. At generation 0, we created up to $n_e = 1,000$ ephemeral (not part of the next generation) offspring, tracking the number of successful mating in n_a attempts, the number of potential mate rejected n_{try} (conditioned on an offspring being produced) and the mean fitness of the resulting offspring (intrinsic $\bar{w}_{e,int}$, extrinsic $\bar{w}_{e,ext}$ and absolute $\bar{w}_e$). For male focal immigrants, this was done using the mate choice rules stated above. For female focal immigrants, it was not possible to conduct this measure within the mate choice framework. Therefore, we used the following approach: a single

Population size	N	10,000
Number of loci	L	500
Number of linkage blocks	n_{LB}	5
Recombination rate between adjacent loci (per individual per generation)	ρ	0.001
Recombination rate between linkage block (per individual per generation)	ρ_{LB}	0.01
Recombination rate between preference locus and first (per individual per generation)	ρ_{MC}	0.1
Number of phenotypic traits	n_t	1
Mutation rate (per individual per base pair per generation)	μ	10^{-6}
Standard deviation of mutational effects	σ_μ	0.05
Standard deviation for mutation at the preference loci	σ_P	0.05
Standard deviation for mutation at the cost loci	σ_C	0
Shape of the FGM, controlling the average epistasis	Q	2
Optimal phenotype at trait j for environment A (and B)	z_j^A (z_j^B)	5 (-5)
Squared distance between perfectly adapted individuals from environment A and B	Δ_{AB}	100
Fitness of an optimal phenotype for environment B (A)in the A (B) environment	Ω^B (Ω^A)	0.1
Probability for a male perfectly adapted to environment A (resp. B) to accepting a female perfectly adapted to environment B (resp. A)	P_c	1
Number of mating before failure for a choosy individual	χ	100
Number of pairwise DMIs	n_D	50
Epistasis due to DMIs	ϵ	0.1
Migration rate	m	0.05

Table 1: List of parameters and their default values

male was selected according to its fitness and tried to mate with the focal female until success or χ attempts happened, with the chance of acceptance to mate and the maximum number of attempts both dictated by the male preference and male cost of preference (this procedure was repeated n_e time, with a potentially different male each time).

Simulations were written using C++ (2011), with the gsl (v2.7, Galassi et al. [2002]) and boost (v1.80) libraries. Analysis was conducted using Mathematica (v12, Wolfram Research, Inc. [2019]), RStudio (v2022.07.2+576) and R (v4.2.0, R Core Team [2022]), with figures generated using the ggplot2 plackage (v3.5.1, Wickham [2016]).

Results

Evolution of strict mate choice or accumulation of genetic incompatibilities depends on migration rate

For each parameter combinations, we performed 10 replicate simulations over $50N$ generations. Of the three barriers investigated, local adaptation yielded the most consistent results, with limited variance both within and across replicates (Fig. 2A). In contrast, both mate choice (Fig. 2B) and the accumulation of genetic incompatibilities (Fig. 2C) displayed more variance between replicates. For the rest of this section, we will mostly refer to the mean effect across replicates, but figures show all simulations to provide a visual representation of the variance.

We did not find any case where the three barriers evolved simultaneously consistently across all 10 replicates (Fig. 3A). There were only 8 cases out of 190 where both local

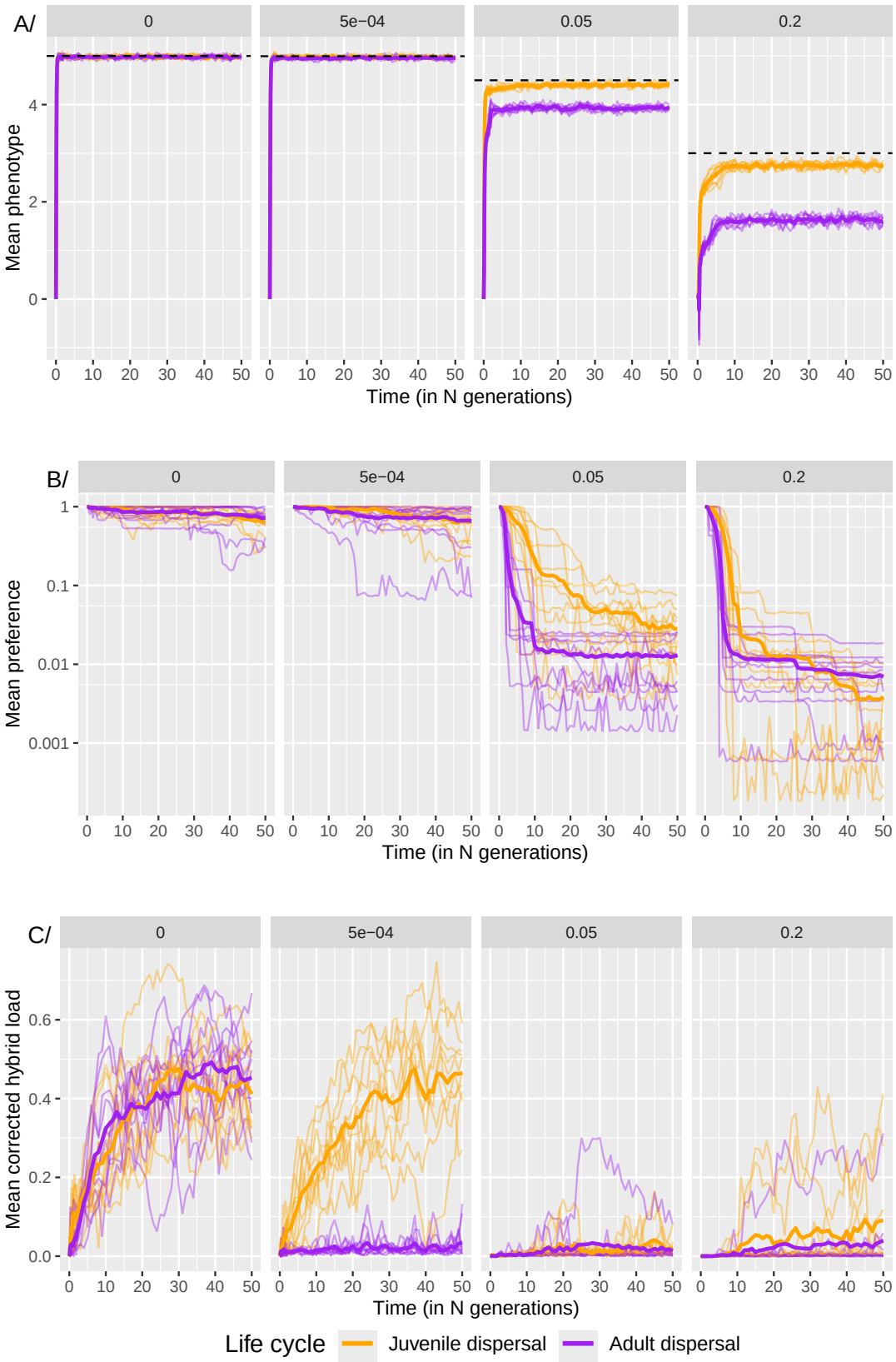

Figure 2: Evolution of the phenotype (panel A), mate choice (B; on a log scale) and hybrid load (C) for different migration rates (given in the facet header) and different life cycles ("juvenile dispersal" in orange and "adult dispersal" in purple). Each thin line corresponds to a different replicate, with the thick line the mean over the 10 replicates. For the top panel, the red dashed line depicts the expected mean phenotypic values, given in equation (5). Relation between the evolved of mate choice and hybrid load are displayed in Figures 3 and S6. The genetic map used was the "default", populations started in the "ancestral" state and other parameters correspond to default values given in Table 1.

adaptation, genetic incompatibilities (corrected hybrid load > 0.2) and strict mate choice (mean preference < 0.1) evolved, but these cases were spread across different parameter combinations. We identified two possible regimes in the evolution of the three mechanisms of reproductive isolation (Fig. 2). Local adaptation was always the first mechanism to establish, but whether it was followed by the build-up of intrinsic postzygotic genetic barriers (accumulation of genetic incompatibilities) or the evolution of prezygotic barriers (strict conspecific mate choice) depended on the migration rate. This is not surprising, as migration rate strongly impacts the effectiveness of both mate choice and genetic incompatibilities.

In the absence of migration, populations evolved toward their phenotypic optimum, with pairs of individuals (within populations) having similar phenotypes, with limited variance at the phenotypic and genetic level. In the limited migration regime, once local adaptation evolved it was followed by accumulation of DMIs, but no evolution of mate choice. This regime of local adaptation, (almost) random mating and (passive) DMIs accumulation across populations extended to low migration rates, with the threshold depending on the life cycle (up to $m < 5 \times 10^{-5}$ for "adult dispersal" and $m < 5 \times 10^{-4}$ for "juvenile dispersal", Fig. 3, S2). Mate choice behaved (almost) neutrally in this context. Indeed, with limited migration, evolving towards strict assortative mating was not advantageous due to limited phenotypic variance, and due to lack of a fitness cost associated with being choosy (as long as males were allowed a second time i.e., $\chi > 1$;Fig. S3).

Because mate choice could range from random mating (preference value of $P_c = 1$) to strict assortative mating, some weak level of preference would still evolve neutrally due to mutational bias (when too close to the extrema values, mutation and drift displaced traits away from the extreme values). Due to isolation, as expected, DMIs accumulated between populations by chance, with purifying selection preventing the accumulation of DMIs within each population. Since incompatible alleles were also involved in local adaptation, they could be under positive or negative selection, depending on the genetic background and environment. In addition, the number of DMIs that could be established between populations was limited, as within each population *de novo* mutations at the interacting loci (i.e., if allele A_1 is fixed, mutation from b_1 to B_1 created a mutational load, see section below) were removed.

In the moderate to high migration regime, once populations became locally adapted, strict mate choice evolved, but there was no accumulation of DMIs. This corresponded to a regime of local adaptation, strict assortative mating and no DMIs. This can be explained by the high intrinsic and extrinsic hybrid costs resulting from mates between individuals adapted to the two environments. This is because when individuals with a phenotype close to the optimum of environment A mated with individuals with a phenotype close to the optimum of environment B, the resulting offspring had both unfit phenotypes (low extrinsic fitness component), and could express potential DMIs (low intrinsic fitness component). Under these conditions, there was a strong selective pressure towards the evolution of assortative mating to avoid these unfit matings. Conversely, under these conditions, migration prevented the accumulation of DMIs. As migration led to gene flow

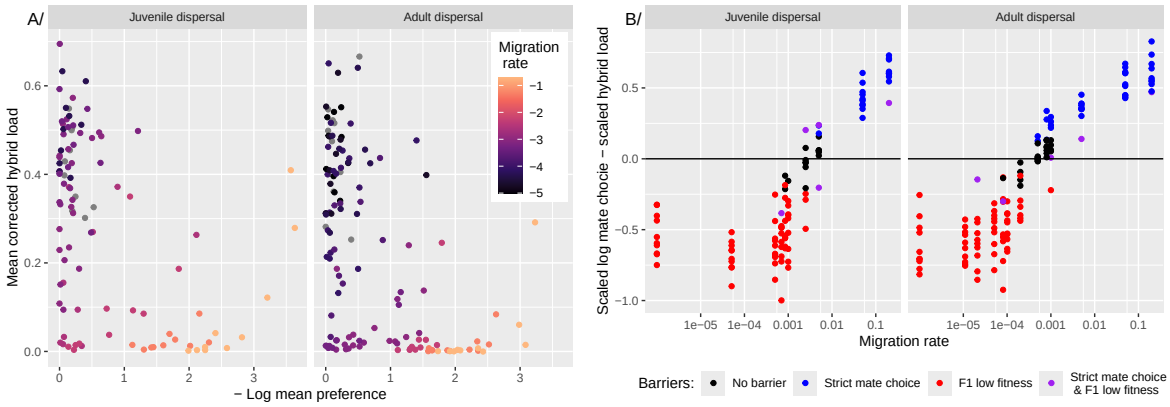

Figure 3: Evolution of strict mate choice and accumulation of DMIs are (mostly) exclusive and depends on the migration rate. A/ Relation between (-log) preference (strict mate choice corresponds to large values) and hybrid load (right panel) at the end of the 50N generations. Color indicates the migration rate (on a log-scale, with $m = 0$ in gray). A detailed figure of this relation per migration rate is given in Figure S6 B/ Relative contrast between scaled log preference (defined between 0 (no mate choice) and 1) and scaled hybrid load (defined 0 (no hybrid load) and 1) as a function of migration rate ($m + 10^{-6}$ on a log-scale) and life cycle. Color indicates whether the reproductive isolation mechanism has evolved (scaled values > 0.25). Other parameters correspond to the default value and are given in Table 1.

between the two populations, selection against derived alleles at each DMI pair would remove them. Thus, despite having (strong) local adaptation and strict mate choice, gene flow was still pervasive enough to trigger selection against DMIs and prevent their accumulation across populations.

Interestingly, we found that despite strong extrinsic and pre-zygotic barriers to gene flow, both in terms of local adaptation (the optimum phenotype of environment A was $\sim$ 10x fitter than an immigrant with the optimum phenotype of environment B, but only 1.78x fitter than the F_1 phenotype), and mate choice (a male at optimum phenotype of A will accept mating with a immigrant female with the optimum phenotype of B in $\sim 1\%$ of the case), DMIs did not accumulate when migration rate was higher than 0.005. The only difference for the case of migration versus limited migration, was that incompatible alleles were expressed in $F_1 s$, creating an intrinsic hybrid load. This indicated that provided there was sufficient gene flow (i.e., through $F_2 s$ and further backcrosses), selection against incompatible alleles involved in DMI prevents their accumulation, even for relatively low migration rates ($m = 0.005$). Indeed, the F_1 hybrid load rarely led to a reduction of more than 10% of the fitness and such values did not remain over long time scale. The presence of gene flow was further confirmed by looking at the distribution of phenotype at the end of the simulations (Fig. S4): we observed patterns indicative of the existence of not only F_1 individuals (with a phenotype close to 0), but also of further backcrosses into each population. We also checked the genome of individuals at the end of the simulations (Fig. S5), further supporting that the low hybrid load was due to the absence of DMIs.

The life-cycle influenced the migration rate at which we saw a transition between the two regimes (Fig. 3B), with a lower migration rate threshold for the "adult-dispersal" life-

cycle. This is expected because immigrating individuals are more likely to contribute to the reproductive pool in the "adult-dispersal" life-cycle, as their fitness depends on their birth environment. Thus, individuals with maladapted phenotypes were more likely to mate in the "adult dispersal" regime, leading to stronger selective pressure to avoid them. Thus, evolution of strict mate choice occurred at lower migration rates for the "adult-dispersal" life-cycle. Interestingly, we did not find a parameter space where both strict mate choice and DMI evolved simultaneously at significant strength (Fig. 3B, no values of m with a majority of purple points). This suggested that selection against incompatibilities was more effective than selection for stricter mate choice. Yet, within a given set of parameters, we observed a correlation between stricter mate choice and higher hybrid load, suggesting a trend for the two mechanisms to coevolve, in particular for moderate to large migration rate (Fig. S6). As a consequence, we would expect that reinforcement (here, we mean the evolution of strict mating due to genetic incompatibilities) would be unlikely under our model. Instead, once mate choice is sufficiently strong, DMIs start to (relatively slowly) accumulate between the population (found for some replicates at high migration $m = 0.2$, Fig. 3).

To further confirm the absence of reinforcement, we investigated secondary contact cases (Fig. 4, with or without initial strict assortative mating). We found that post-zygotic barriers due to DMIs erode over time, but with two different regimes. When secondary contact occurs between populations without mate choice ("SC"), we found that the evolution of mean preference towards strict mate choice was similar to when two populations diverged from an ancestral population ("Ancestral"), and the hybrid load due to DMIs converged to similarly low values in both scenarios (Fig. 4 left vs mid panels). In contrast, if secondary contact occurred between populations already with DMIs and strict mate choice ("SC + initial mate choice"), we found a distinct dynamics, with evolution of hybrid load towards high equilibrium values (irrespective of the migration rate), and maintenance of very strict mate choice for scenarios with high migration ($m > 0.05$). Importantly, for high migration ($m = 0.2$, Fig. 4 right panels), the mate choice preference for "SC + initial mate choice" converged to lower values (stricter mate choice) than for the "SC" scenario, suggesting that if preference evolved towards very low values (i.e., very strict mate choice) it could prevent gene flow and DMIs could then accumulate. The post-zygotic barrier due to DMIs eroded over time under all considered scenarios, either towards the hybrid load equilibrium (see next section) if strict assortative was already present ("SC + initial mate choice"), or until it was mostly gone. We note that erosion of the post-zygotic barrier due to DMIs happened relatively slowly over time. This is a remarkable point since the absence of an observed hybrid load when assortative mating evolved to became strict could be explained by the formation of a subpopulation of (unfit) individuals from the alternative environment. This substructure could lead to preferential matings of such individuals with focal immigrant individuals used to measure hybrid load between populations (i.e., matings not expressing the DMIs). Yet, the slow decline indicates that DMIs are purged rather than not expressed due to substructure. Overall, this indicates that reinforcement (as defined above) is not possible under the investigated scenarios, due to the effective purging of DMIs when there is enough gene

flow ($m \geq 0.005$).

No accumulation of DMIs with migration and lack of snowball effect, irrespective of the network of DMI interactions

As illustrated on Figure 2C and pointed out in the previous section, the intrinsic hybrid load was either absent or evolved towards some non-zero equilibrium value. In the later case, the hybrid load matched what was observed in the absence of migration (and therefore absence of gene flow) and could be obtained starting from an ancestral population or a secondary contact scenario. Such value of the hybrid load was equivalent to having five DMI pairs being differently fixed between the two populations. This result indicated that in scenarios where DMIs do accumulate, only about 10-15% of all possible DMIs between the two populations were differently fixed (in terms of equivalent number of DMIs; Fig. S8). This pattern could be partially explained by the design of the genetic architecture and in particular, by the fact that forward and backward mutation rates (at each biallelic locus) were considered equal.

To be better understand why only 10 to 15% of the DMIs were differentially fixed, we considered a simple deterministic model describing how DMIs would accumulate between populations in the limit of weak mutation and strong selection (epistasis). In each population, at equilibrium and assuming a low mutation rate, about half of the loci would be fixed for the derived allele (i.e., A_i or B_j) and the other half for the ancestral allele (i.e., a_i or b_j). In addition, for each DMI, the derived alleles (A_i or B_i) could only arise in the ancestral background, whereas the ancestral alleles could arise in any background. Under these conditions, the population had three possible states: fixed for a_ib_i, A_ib_i and a_iB_i. The expected probability of having A_i fixed in the first population at a given time-point was $P(\text{A in } p_1) = \frac{1}{3}$ (more generally, if we defined ν as the back mutation rate, from A_i to a_i, this probability was given by $P(\text{A in } p_1) = \frac{\mu}{2\mu+\nu}$). Therefore, the probability of having a DMI between the two isolated populations was $\frac{2}{9}$ (since $P(\text{DMI}) = P(\text{A in } p_1)P(\text{B in } p_2) + P(\text{A in } p_2)P(\text{B in } p_1) = 2 * P(\text{A in } p_1)^2$ or $2\left(\frac{\mu}{2*\mu+\nu}\right)^2$, since $P(\text{A in } p_1) = P(\text{A in } p_2) = P(\text{B in } p_1) = P(\text{B in } p_2)$. In the absence of back mutation, this probability became $P(\text{A in } p_1) = 0.5$.

We observed a lower expected number of DMIs in the simulations than predicted by the weak mutation strong selection approximation. However, further exploration revealed that neither a change in epistasis strength to either $\epsilon = 0.9$ or 0.01 (Fig. S9) or a change in the genetic map ("default", "network" or "optimized" maps) affected the expected number of DMIs. Finally, more important than changes in effective size, what determined the proportion of equivalently fixed DMIs was the scaled mutation rate (i.e., $\theta = 4N\mu$, Fig. S10). These results suggested that, in large populations or when mutation rate is high, multiple mutations might segregate at the same time. In particular, if both A_i and B_i segregated in a given population, selection against those incompatible alleles would prevent their fixation.

We further investigated whether the DMI architecture could explain the lack of a snowball effect, which predicted an increasing number of DMIs through time. The snowball

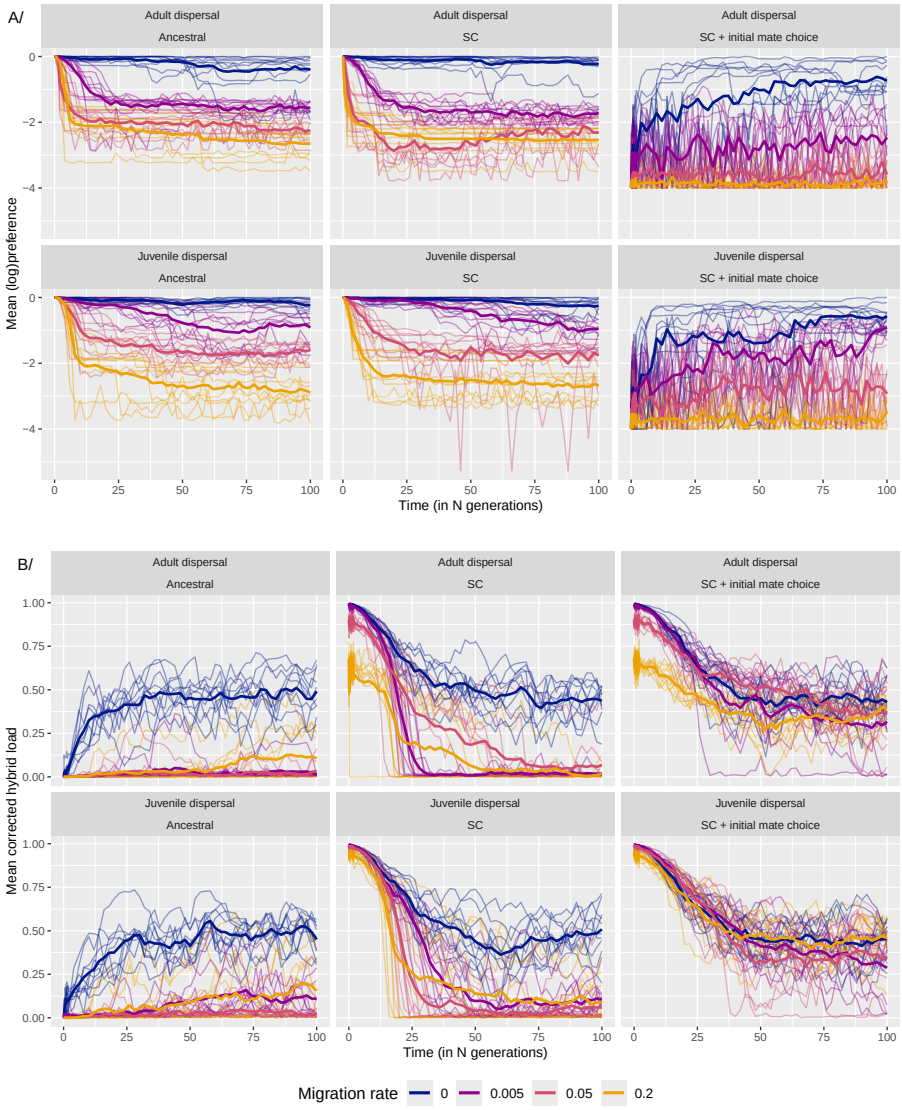

Figure 4: Evolution of mate choice (panel A) and F1 hybrid load (B) for different starting conditions: the "Ancestral" (default) scenario, the secondary contact ("SC") scenario and the secondary contact with initial mate choice scenario ("SC + initial mate choice"). Different migration rate are given indicated by different colors, with each thin lines corresponding to a different replicate and the thick line the mean over the 10 replicates. The final states and link between mate choice and hybrid load are displayed in Figure S7. Other parameters correspond to the default value and are given in Table 1

effect predicted by Orr [1995] is expected to be mediated by whether incompatible loci are involved in a single or multiple DMIs Guerrero et al. [2017]. We thus considered a "network" architecture where each locus could be involved in more than one pair of DMIs (not associated with the phenotype), and an "optimized" architecture where for each DMI pair the A_j allele at locus A had a positive phenotypic effect (favored in environment A) and allele B_j at locus B had a negative phenotypic effect (favored in environment B). Note that given the same number of DMI pairs, the "network" architecture has less loci involved in incompatibilities. We found evolution towards lower hybrid load values for the "network" than the "default" and "optimized" ones, suggesting that the lower the number of underlying DMI loci, the lower the number of equivalent fixed DMIs (Fig. 5B). This is due to the larger number of interacting loci that increases the selective pressure, preventing the fixation of derived alleles, even in the absence of gene flow. Indeed, under the "network" architecture, we did not find an accumulation of DMIs, even in scenarios without migration. In contrast, for the "optimized" architecture, despite the fact that DMIs were fully correlated with the phenotypic effect of alleles, the proportion of equivalent fixed DMI evolved towards values similar to the "default" architecture (approximately 0.15). Importantly, with migration, accumulation of DMIs was strongly limited in all investigated DMI architectures.

Finally, we found that mate choice and DMIs could not maintain population differentiation in absence of local adaptation in secondary contact models (Fig. S11). Interestingly, at high migration rate, the lack of differentiation was not caused by a collapse of reproductive isolation, but through initial ecological exclusion of one of the two populations (Fig. S11). Once one of the populations was lost, the role of mate choice and genetic incompatibilities became irrelevant and individuals in two environments behaved as a single population. In contrast, at low migration rates ($m \leq 0.0005$), we found that phenotypic differentiation converged to zero at similar rates for cases with and without migration, indicating that this is due to mutational bias (see section below). However, DMIs were actively purged with migration, as the hybrid load converged to zero (Fig. S11).

Shape of ecological fitness landscape has reduced impact in the evolution of reproductive isolation barriers

We investigated whether the number of dimensions (i.e., number of traits) of the phenotypic space affected the evolution of the different mechanisms of reproductive isolation. As the number of phenotypic traits increased, the distance between the optimal phenotypes increased, therefore we set the position of the optima for each phenotypic trait such that the distance between optima at the two populations was constant, irrespective of the number of dimensions (for $n_t = 3$, the optimum for environment A is $(2.887, 2.887, 2.887)$ and $(-2.887, -2.887, -2.887)$ for B). Interestingly, we did not observe any significant changes in the evolution of reproductive isolation for all three mechanisms at different migration rates (Fig. S12).

We also investigated the shape of the fitness landscape. In principle, this parameter could have a large impact on the evolution of reproductive isolation, as it will impact the

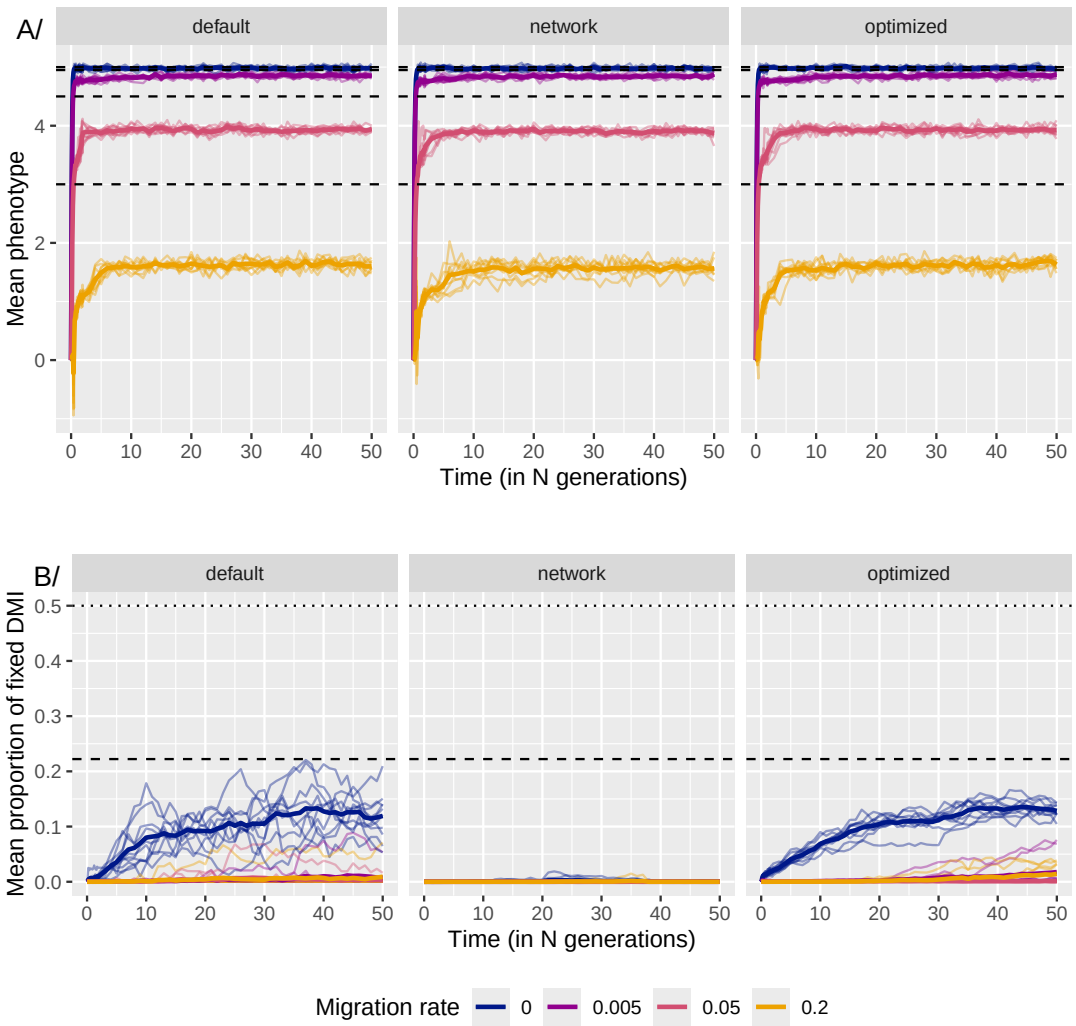

Figure 5: Evolution of the phenotype (panel A) and number of equivalent fixed DMIs (B; computed according to equation (6)) for different DMIs architecture (given in the facet header) and different migration rates. Each thin line corresponds to a different replicate, with the thick line the mean over the 10 replicates. Color corresponds to different migration rates: $m = 0$ in blue, $m = 0.005$ in purple, $m = 0.05$ in pink and $m = 0.2$ in orange. For the A panel, the black dashed line depicts the expected mean phenotypic values for each migration rate, given in equation (5). The "default" network corresponds to the 100 loci out of 500 forming 50 pairs of DMIs, the "network" one to 50 loci out of 800 forming 200 pairs of DMIs, and the "optimized" one to 500 loci out of 500 forming 250 pairs of DMIs, with always an allele increasing the phenotypic value interacting with one decreasing it (see methods). Other parameters correspond to the default value and are given in Table 1

fitness of phenotypes in the vicinity of the optimum. When mean epistasis between mutations was positive ($Q < 2$), the fitness landscape became sharper, with small deviations in phenotype around the optimum leading to strong reductions in fitness. Oppositely, when mean epistasis was negative ($Q > 2$), the fitness landscape around the optimum became flatter (i.e., an individual with a phenotype away from the optimum can still have a high fitness). In addition, since the fitness of F_1 hybrid individuals between parents at the optima of the two environment (and therefore having an intermediate phenotype) is given by w_{alt}^{2-Q}, increasing Q weakens the role of local adaptation in reproductive isolation. As illustrated in Figure S13 (for $m = 0.05$), the mean phenotype of the population sat further away from the optimum as Q increased. Yet, strict mate choice evolved regardless of the mean epistasis in the fitness landscape. Finally, accumulation of DMIs was only observed for $Q < 2$ and only in some replicates.

To assess if our conclusions were affected by the inherent mutational bias in our model towards the phenotype value of zero for all traits (i.e., the same as the phenotypic value of F_1 hybrids for cases with symmetric optima), we repeated the simulations under models with asymmetric optima at the two populations. In the presence of local adaptation, we did not detect any changes due to the asymmetry (Fig. S14). The deviation of the mean phenotype from the optimum was the same regardless of the relative position of the ancestral phenotype and the environmental optima, indicating that the effect of selection and migration overshadow the effect of mutational bias.

Gene flow mostly caused by the choosy sex

Due to the implementation of the mate choice scheme, we expected that a choosy individual (here a male) would have a higher chance to leave offspring after moving to the alternative environment than a non-choosy one (a female), all things being equal. This arose because a male (choosy individual), once selected as a potential mate based on extrinsic fitness (based on phenotypic value), had χ attempts to accept one mate, while the female, once chosen as a potential mate, had only one chance to be accepted for mating. Assuming all individuals were at the optima in each environment, this difference could easily be quantified, considering that the resident population had size $2N$ with 50:50 sex-ratio, all resident individuals had fitness w_r, and that the focal immigrant individual had fitness w_f. Under these assumptions, the probability that this focal individual left offspring in the next generation became for males:

$$P_m(desc.) = 1 - \left(\frac{(N-1)w_r}{w_f(1-(1-P_c)^{\chi}) + (N-1)w_r}\right)^{2N} \xrightarrow[N\to\infty]{} 1 - e^{2((1-P_c)^{\chi}-1)\frac{w_f}{w_r}} \quad (7)$$

and for females :

$$P_f(desc.) = 1 - \left(\frac{(N-1)w_r}{P_c w_f + (N-1)w_r}\right)^{2N} \xrightarrow[N\to\infty]{} 1 - e^{-2P_c\frac{w_f}{w_r}} \quad (8)$$

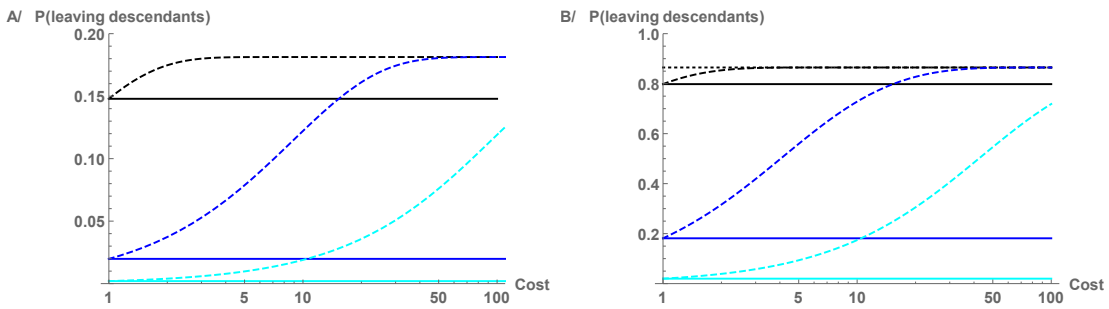

Figure 6: Probability of a focal immigrant individual to leave at least one descendant in the population in the next generation as a function of the cost of choosiness, in the absence (panel A, $w_f = w_r$) and presence (B, $w_f = 0.1$, $w_r = 1.0$) of local adaptation. We represent this probability a monomorphic resident population for males (dashed line; eq. 7) and females (eq. 8 solid line) and difference preference strength ($\pi = 0.8$ in black, $\pi = 0.1$ in blue and $\pi = 0.01$ in cyan). The probability for the neutral case is given by the dotted dashed line on panel A and is not displayed on panel B due to scale. Other parameters are $N = 1000$ and $w_r = 1$.

where P_c was the probability of a male accepting a mating with a female with a different phenotype and χ the number of mating attempts a male did before being discarded. Details of the calculations are given in the Appendix. For reference, under complete neutrality and random mating, the probability that the focal individual left descendant was $P(desc.) = 1 - (\frac{N-1}{N})^{2N} \xrightarrow[N\to\infty]{} 1 - \frac{1}{e^2} \approx 0.865$ (matching our simulations Fig. S15). These equations were illustrated in Figure 6 (and matched simulations with the active barriers, Fig. S16). As expected, male mate choice meant that immigrating females were far less likely to leave immediate descendants than their male counterpart. In addition, as the cost of being choosy weakened (χ increased), the effect of the strength of preference vanishes for males, while it remained the same for female. Furthermore, for strong preference ($P_c \to 0$), the chance of males having immediate descendants were χ times higher than females; this ratio would decrease as preference increased, with all differences vanishing in the absence of preference ($P_c = 1$). The difference between sexes also vanished when the cost of choosiness was extremely high $\chi = 1$. Overall, this indicated that under strong preference ($P_c \to 0$) and weak preference cost ($\chi \to \infty$), introgression is expected to be sex-specific, with the choosy sex being responsible for most of the gene flow between populations.

Local adaptation as the main barrier to gene flow in scenarios with migration

To quantify the barriers to gene flow we computed the mean sojourn times of a neutral allele from a focal immigrant individual. This mean sojourn time is a proxy of the barrier strength, with lower values indicating stronger barriers against gene flow, as alleles from the focal immigrant individual were lost more quickly. This was measured in isolation, to prevent back migration to the source population. As highlighted in the previous section, male and female must be analyzed separately.

Regardless of the migration rate, none of the mechanisms of reproductive isolation increased the mean sojourn time of an unlinked marker compared to the expected time

under neutrality (Fig. 7, panel A for "adult dispersal" and B for "juvenile dispersal"). Furthermore, the combination of barriers was not directly antagonistic in the sense that the joint effect of barriers always decreased the mean sojourn time compared to the effect of each barrier individually. This result was qualitatively the same if we considered a linked marker instead (Fig. S17) or used the harmonic (Fig. S18) or geometric (Fig. S19) means. This was true regardless of the sex of the focal individual, although the choosy sex (males) had longer sojourn times when mate choice was involved ("MC", "LC+MC", "MC+DMI", "all"), consistent with higher gene flow when immigrants were of the choosy sex (see previous section).

Across all combinations of barriers and life cycles, the shortest sojourn times were similar irrespective of the migration rate, indicating that reproductive isolation barriers evolved to similar levels with and without migration. Irrespective of the life cycle and migration rates, cases where all barriers were active ("all") showed shorter sojourn times, i.e., stronger barriers. Yet, cases with local adaptation ("LA"), local adaptation and mate choice ("LA+MC") and local adaptation and genetic incompatibilities ("LA+DMI") often led to similar sojourn times to those found by combining all barriers ("all"), although this depended on the migration rates and life-cycle. In agreement with purging of DMIs, when migration $m \geq 0.05$, the sojourn times when only the DMI barrier was active ("DMI") were similar to neutral, irrespective of life cycle and sex of focal immigrant individual. In contrast, for scenarios without gene flow, DMIs were a more efficient barrier than with migration, leading to a reduction of sojourn times of approximately 0.40.

Since the sojourn time was measured in isolation, the life cycle only affected the process at generation 0, determining whether the fitness of the focal immigrant individual was updated or not in the resident population. Interestingly, across most scenarios we observed differences depending on the life cycle in the reduction of gene flow (compared to neutrality) of each barrier or combination of barriers (Fig. 7). First, for scenarios with local adaptation ("LA", "LA+MC", "LA+DMI", "all"), the reduction in sojourn time was not as strong with the "adult dispersal" life cycle, i.e., local adaptation was a much stronger barrier to gene flow for "juvenile dispersal" scenarios. Second, for the "juvenile-dispersal" cycle, when all mechanisms were active ("all"), the reduction in sojourn time was mostly independent of the migration rate and of the sex of the focal immigrant individual, since focal individuals failed to reproduce in the first generation. For the "adult-dispersal" life cycle, this was no longer the case, as sojourn times depended on migration rate and sex of focal individual (especially striking for $m = 0.005$). The evolution of stronger barriers at intermediate migration rates in the "adult-dispersal" life cycle was due to a balance between two mechanisms: increasing migration triggered the evolution of mate choice, but also ensured that a small proportion of individuals with the alternative phenotype would be in the resident population at each generation. Since in the "adult-dispersal" life-cycle, the focal immigrant individual's extrinsic fitness was the same in the two populations, a migration $m = 0.005$ (Fig. S2) was kind of a Goldilocks parameter zone, where strict assortative mating evolved, but without having too many individuals with the alternative phenotype in the resident environment. Therefore, under the "adult dispersal" life cycle, strong barriers to gene flow leading to speciation would be predicted only for scenarios

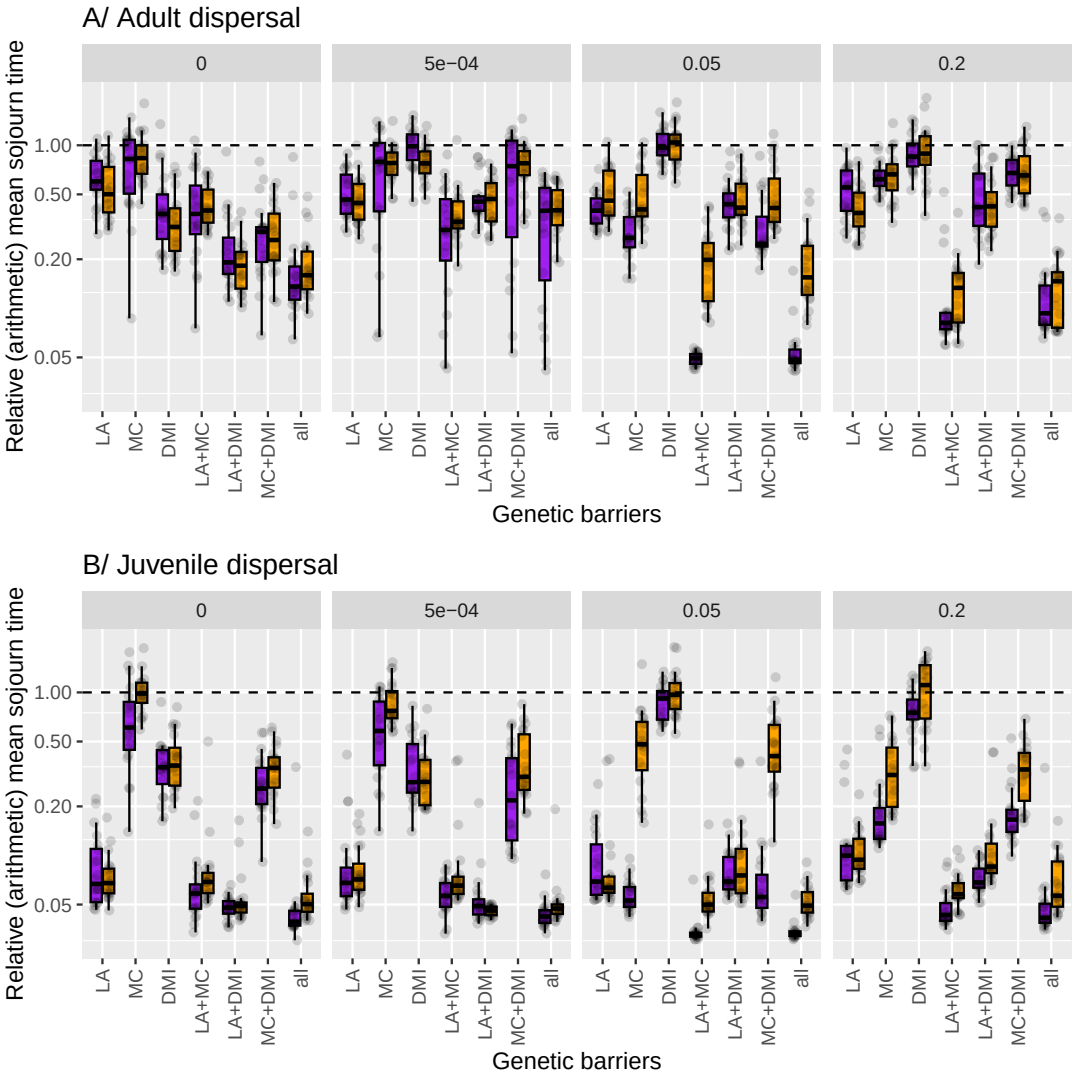

Figure 7: Relative sojourn time of a freely recombining marker introduced by an immigrating individual from the alternative population with different reproductive barriers for males (orange) and females (purple). The sojourn time is measured relative to either the neutral cases with the normalizing factor being the mean sojourn time over replicates. Panel A corresponds to the "adult dispersal" life cycle and panel B to the "juvenile dispersal" one. Additional migration rates are displayed in Figure S20. The corresponding figures for the absolute times (Fig. S21), for the effect of a barrier, in the presence of another (Fig. S22 and S23) and for sojourn time of a linked marker (Fig. S17) are given in Supplement. Other parameters correspond to the default values.

with the "right" amount of migration between populations.

A high variance (overdispersion) for the mean sojourn time under some scenarios was observed, mainly for barriers involving mate choice ("MC", "LA+MC", "MC+DMI", "all") and for intermediate migration rates ($m = 0.0005$ and $m = 0.05$). This overdispersion reflects stochasticity across replicates, since in some replicates, populations evolved strong mate choice, while not in others.

In addition, a couple of important details are seen in Figure 7. First, for $m = 0$, mate choice reduced the sojourn time of a marker introduced by females by approximately 0.80, due to some level of assortative mating evolving because of mutational bias at the preference locus. Males were not affected as the cost of mate choice (χ) was really low (see above). Second, for $m = 0.2$, mate choice was less effective at reducing the relative sojourn time than for $m = 0.05$, despite them having evolved similar preferences (for both life cycles, Fig. 7). This is due to a sufficient proportion of males with the alternative phenotype in the resident population when $m = 0.2$ (in particular for the "adult dispersal" life cycle, Fig. S24), with whom focal immigrant females can reproduce. Therefore, the sojourn time is higher at $m = 0.2$ than $m = 0.05$ since the barrier is much weaker for the sons of the focal female than for herself.

We also investigated the sojourn time of neutral alleles in scenarios of secondary contact with preexisting preference with fixed levels of the strength of each barrier, thus removing the stochastic variation in evolution of barriers across runs. In the "juvenile dispersal" life cycle, local adaption remained the stronger barrier to gene flow, i.e., shorter sojourn times for "LA" (Fig. S25), with addition of mate choice ("LA+MC") further reducing the sojourn times (but only in females, and when each barrier was relatively weak). Genetic incompatibilities ("DMI") were the weakest of all barriers, providing shorter reductions in sojourn time than the other mechanisms, but equivalent to mate choice ("MC") in the choosy sex. For the "adult dispersal" life cycle, local adaptation ("LA") was one of the weakest barriers. Surprisingly, although for the "juvenile dispersal" life cycle, genetic incompatibilities ("DMI") were one of the weakest barriers, for the "adult dispersal" life cycle, "DMI" provided the strongest barrier (i.e., shorter sojourn times), even stronger than local adaptation and with no sex-specific effects.

Discussion

We investigated how the interaction of different mechanisms of reproductive isolation impact divergence and gene flow between populations. In scenarios with migration, local adaptation was unsurprisingly always the first barrier to evolve, with either accumulation of genetic incompatibilities (DMIs) or the evolution of strict assortative mating evolving afterwards. The evolution of the last two barriers were somewhat mutually exclusive: whether strict mate choice or accumulation of DMIs evolved was dictated by migration rate, and to a lesser extent by the life cycle. Surprisingly, even in cases where local adaptation and strict mate choice evolved, DMIs acted as a very weak barrier, consistent with purging of genetic incompatibilities provided there is sufficient gene flow. Thus, reduction in gene flow (strength of reproductive isolation) was mostly a consequence of

the action of local adaptation, with mate choice being also effective. Importantly, we modeled here mate choice as male choice but the model was fully symmetric and all genetic elements were autosomal. Therefore, all results applied to both male and female mate choice.

Local adaption at the center of speciation with gene flow

Under our model accounting for interaction of three barriers, local adaptation due to ecological divergence played a major role in the evolution of reproductive isolation as it was always the first barrier to evolve. Indeed, the necessary aspect of local adaptation for parapatric speciation had already been pointed out previously Agrawal et al. [2011], Bank et al. [2012]. In addition, in the "juvenile dispersal" life cycle, it was also the main mechanism responsible for the reduction of sojourn times of neutral introgressed alleles, a proxy for the level of gene flow between populations. The importance of ecological speciation has received a lot of attention (as detailed in Nosil [2012] - but see Anderson and Weir [2022]). Our results are in agreement with the role of local adaptation in ecological and parapatric speciation scenarios, but bring an important nuance. Indeed, when considering the "adult dispersal" life cycle, local adaptation drives the evolution of phenotypic differentiation between the two populations, and mate choice or accumulation of DMI only evolve once populations become locally adapted. However, provided that populations evolved strict mate choice or accumulated DMIs, the strength of these two barriers becomes similar to local adaptation, as we found for secondary contact models from fixed population states (Fig. S26 versus S25). For both the evolution from an ancestral population or from secondary contact from fixed population states, the relative strength of local adaptation depended on the life cycle, being stronger for the "juvenile dispersal". Since the only difference between the two life cycles is whether the fitness of an immigrant individual is affected or not by the new environment, this indicates that local adaptation is mostly effective at generation one. This result highlights the brittle nature of local adaptation/ecological differentiation as a genetic barrier to gene flow Blanckaert et al. [2020].

Evolution of mate choice when there is sufficient migration

Mate choice was observed to evolve only in the presence of local adaptation and non negligible migration. Therefore it always evolved as reinforcement (sensu Butlin and Smadja [2018]). In secondary contact scenarios, mate choice was not able to self-maintain without the presence of ecological differentiation. However, the loss of mate choice as an effective reproductive barrier in the secondary contact scenarios was due to competitive exclusion of one of the two species, and not a collapse of the reproductive barrier itself. Mate choice was also effective at limiting gene flow between species, but only for the non choosy sex and when a strong preference evolved. This is in line with what was observed by Irwin [2020] in the role of assortative mating to maintain a narrow hybrid zone between populations. As in Irwin [2020], our results partly stem from modeling

preference using a Gaussian function and a matching phenotype approach. In particular, this implies that mating of F_1s with identical phenotypes is very likely, and F_1s are also likely to be accepted as potential partners by resident individuals. Our approach to model mate choice creates a "now or never" dichotomy for reducing gene flow, which is seen in differences between the choosy and non-choosy sex. Having the preference function more akin to a step function instead of a Gaussian could mitigate the effects described here and generate stronger barriers to gene flow. The suitability of each approach depends on whether the trait used for mate choice is qualitative or quantitative. In the later case, a continuous smooth function is a more likely representation of the biological process involved in mate choice (especially since we are not considering the environmental variance that could contribute to the phenotype and would flatten the preference function). Finally, importantly, we never observed the evolution of mate choice due to the presence of genetic incompatibilities, as the two mechanisms seem to exist in different parameter spaces. We only found accumulation of DMIs with limited migration, where the benefit to evolving mate choice is negligible. This is likely because the selection to purge DMIs is more sensitive to migration than selection to trigger the evolution of mate choice.

Genetic incompatibilities do not accumulate in locally adapted populations with strict mate choice when there is sufficient gene flow

Two types of incompatibilities were modeled in our approach: extrinsic and intrinsic ones. Indeed, Fisher Geometric Model naturally generates epistasis between mutations due to the quadratic relationship between phenotype and fitness. Interestingly, most of the empirical patterns can be mimicked within the model framework Fraïsse et al. [2016]. However, such approach makes it extremely challenging to characterize the effect of genetic incompatibilities, as they are environment-dependent and depend on the whole genetic background. Therefore, our model also included intrinsic incompatibilities using the classical Bateson-Dobzhansky-Muller model Bateson [1909], Dobzhansky [1936], Muller [1942]. Our study reveals a duality in the effect of these intrinsic genetic incompatibilities in the speciation process. Accumulation of DMIs is relatively constrained, only happening when migration was absent or limited. This effect was further amplified when forward and backward mutation rates were symmetric, as the within-population mutational load generated by the possible DMIs tend to select against the fixation of any of the incompatible alleles (Fig S10), even in the absence of migration.

As soon as migration was non negligible, genetic incompatibilities were purged from the system, despite the evolution of phenotypic differentiation and strong mate choice. Interestingly, this pattern has been recurrently observed in nature. For example, cichlid fish species maintain their integrity despite coexisting in at least parapatry, and in the presence of mate choice, yet hybrid breakdown remained limited Stelkens et al. [2015] and DMIs accumulation happened at a much slower rate than speciation Stelkens et al. [2010]. Such pattern also matches the example of *Anopheles gambiae* and *A. coluzzi*. As summarized in Mallet and Mullen [2022], the two species display ecological differences and strong assortative mating, but are sympatric and apart from a few islands of differentiation, 99%

of the genome is homogenized, and therefore not propitious to the accumulation of DMIs Turner et al. [2005], Lee et al. [2013]. The lack of characterized DMIs is partly due to the challenge associated with their detection Nouhaud et al. [2020], Blanckaert and Payseur [2021], but our model here also highlights that low levels of gene flow between populations is sufficient to purge DMIs. Overall, under our model, genetic incompatibilities seems to arise more as a consequence of reproductive isolation, and not as an active mechanism to prevent gene flow, matching the predictions from the two-locus model Bank et al. [2012]. An important aspect to these result is that we are assuming constant population size and therefore soft selection Wallace [1968, 1975] - low fitness phenotypes therefore have a higher chance of leaving offspring when frequent enough compared to a hard selection scenario. While extinction of either or both population is a non negligible risk, demographics and evolutionary mechanisms have been shown to act in opposite direction with regards to the hybrid speciation process Blanckaert et al. [2023] and therefore DMIs may play a more important role when considering different assumptions - such as allowing the populations to be affected by bottlenecks.

Despite this, multiple DMIs have been reported across many organisms Presgraves [2003], Kao et al. [2010], Corbett-Detig et al. [2013], Powell et al. [2020], with the underlying genetic mechanisms responsible having been characterized (reviewed in Table 1 of Kitano and Okude [2024], including 8 cases of dominant DMIs). Potentially, all of these DMIs may have formed in isolation, as described in the conceptual definition of the DMI model Coyne and Orr [2004] whether due to temporal, geographical or prezygotic barrier. Yet, gene flow between these species has been identified (for example Powell et al. [2020] for the *Xiphophorus malinche* and *X. birchmanni*). Furthermore, Dettman et al. [2007] was able to observed the evolution of DMIs as a by-product of divergent selection between isolated populations of *Saccharomyces cerevisiae*, matching our results. Whether such mechanism could happen with ongoing gene flow remain largely unexplored Kulmuni and Westram [2017]. Alternatively, it is is possible that DMIs accumulation happens at a much slower rate, as discussed for cichlids Stelkens et al. [2010], that was no captured with our design here. Finally, DMIs are often thought as fixed differences between species. Yet as argued by Cutter [2012], this may not be the case, and many genetic incompatibilities may be segregating in the populations - a feature we recovered in our simulations (Fig. S27).

Conclusions

We explicitly modeled the interaction between local adaptation, genetic incompatibilities (DMIs) and mate choice, by considering that the same loci could be involved in the phenotype (under selection in the two environments) and associated with DMIs, and that mate choice depended on phenotype and that preference was due to a modifier locus linked to sites determining the phenotype. Surprising, despite a common genetic basis, we found that long term evolution from an ancestral population did not lead to the joint build up of the three reproductive isolation barriers, i.e., locally adapted populations with strict mate choice and accumulation of DMIs. Instead, we found that populations evolved towards

two regimes: either local adaptation and accumulation of DMIs with weak mate choice, or local adaptation with strict mate choice and no DMIs. The migration rate determined which regime was possible, with higher migration rates favoring the latter. It is thus not surprising that we recover results similar to the ones that have been derived for when two of the three reproductive isolation mechanisms interact. This implies that under the genetic architecture with polygenic adaptation and multiple DMIs considered here, it is unlikely that populations evolve towards build up of strong pre- and post-zygotic barriers when there is gene flow. How loci involved in local adaptation and DMIs are distributed along the genome, and whether these different regimes lead to genetic signatures that can be detected based on polymorphism population genomics data remains to be further studied.

1 References

References

A. Agrawal, J. Feder, and P. Nosil. Ecological divergence and the origins of intrinsic postmating isolation with gene flow. *International Journal of Ecology*, 2011(1):435357, 2011.

S. Aguillon, S. Haase Cox, Q. Langdon, T. Gunn, J. Baczenas, S. Banerjee, A. Donny, B. Moran, C. Gutiérrez-Rodríguez, O. Ríos-Cárdenas, et al. Pervasive gene flow despite strong and varied reproductive barriers in swordtails. *bioRxiv*, 2024.

S. Anderson and J. Weir. The role of divergent ecological adaptation during allopatric speciation in vertebrates. *Science*, 378(6625):1214–1218, 2022.

T. Aubier, M. Kopp, I. Linn, O. Puebla, M. Rafajlović, and M. Servedio. Negative coupling: the coincidence of premating isolating barriers can reduce reproductive isolation. *Cold Spring Harbor Perspectives in Biology*, page a041435, 2024.

C. Bank, R. Bürger, and J. Hermisson. The limits to parapatric speciation: Dobzhansky–Muller incompatibilities in a continent–island model. *Genetics*, 191(3):845–863, 2012.

T. Barraclough. Does selection favour the maintenance of porous species boundaries? *Journal of Evolutionary Biology*, page voae030, 2024.

N. Barton. Multilocus Clines. *Evolution*, 37(3):454, 1983.

W. Bateson. Heredity and variation in modern lights. *Darwin and modern science*, pages 85–101, 1909.

A. Blanckaert and J. Hermisson. The limits to parapatric speciation II: Strengthening a preexisting genetic barrier to gene flow in parapatry. *Genetics*, 209(1):241–254, 2018.

A. Blanckaert and B. Payseur. Finding hybrid incompatibilities using genome sequences from hybrid populations. *Molecular biology and evolution*, 38(10):4616–4627, 2021.

A. Blanckaert, C. Bank, and J. Hermisson. The limits to parapatric speciation 3: Evolution of strong reproductive isolation in presence of gene flow despite limited ecological differentiation. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 375(1806), 2020.

A. Blanckaert, V. Sriram, and C. Bank. In search of the goldilocks zone for hybrid speciation ii: hard times for hybrid speciation? *Evolution*, 77(10):2162–2172, 2023.

D. Bolnick, A. Hund, P. Nosil, F. Peng, M. Ravinet, S. Stankowski, S. Subramanian, J. Wolf, and R. Yukilevich. A multivariate view of the speciation continuum. *Evolution*, 77(1):318–328, 2023.

R. Butlin and C. Smadja. Coupling, reinforcement, and speciation. *The American Naturalist*, 191(2):155–172, 2018.

R. Butlin, J. Galindo, and J. Grahame. Sympatric, parapatric or allopatric: the most important way to classify speciation? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1506):2997–3007, 2008.

D. Castillo, M. Burger, C. Lively, and L. Delph. Experimental evolution: assortative mating and sexual selection, independent of local adaptation, lead to reproductive isolation in the nematode caenorhabditis remanei. *Evolution*, 69(12):3141–3155, 2015.

R. Corbett-Detig, J. Zhou, A. Clark, D. Hartl, and J. Ayroles. Genetic incompatibilities are widespread within species. *Nature*, 504(7478):135–137, 2013.

J. Coyne and H. Orr. *Speciation*. Sinauer Associates Sunderland, MA, 2004.

M. Cruz, S. Magalhães, M. Bakirdöven, and F. Zélé. Wolbachia strengthens the match between pre-mating and early post-mating isolation in spider mites. *bioRxiv*, pages 2024–05, 2024.

A. Cutter. The polymorphic prelude to Bateson–Dobzhansky–Muller incompatibilities. *Trends in ecology & evolution*, 27(4):209–218, 2012.

C. Darwin. *The origin of species*. London: Murray, 1859.

J. Dettman, C. Sirjusingh, L. Kohn, and J. Anderson. Incipient speciation by divergent adaptation and antagonistic epistasis in yeast. *Nature*, 447(7144):585–588, 2007.

T. Dobzhansky. Studies on hybrid sterility. II. Localization of sterility factors in *Drosophila pseudoobscura* hybrids. *Genetics*, 21(2):113, 1936.

E. Dopman, K. Shaw, M. Servedio, R. Butlin, and C. Smadja. Coupling of barriers to gene exchange: causes and consequences. *Cold Spring Harbor Perspectives in Biology*, page a041432, 2024.

M. Feldman, S. Otto, and F. Christiansen. Population genetic perspectives on the evolution of recombination. *Annual review of genetics*, 30(1):261–295, 1996.

R. A. Fisher. *The genetical theory of natural selection.* Oxford University Press, Oxford, 1930.

C. Fraïsse, P. A. Gunnarsson, D. Roze, N. Bierne, and J. J. Welch. The genetics of speciation: Insights from fisher's geometric model. *Evolution*, 2016.

Mark Galassi, Jim Davies, James Theiler, Brian Gough, Gerard Jungman, Patrick Alken, Michael Booth, Fabrice Rossi, and Rhys Ulerich. *GNU scientific library.* Network Theory Limited Godalming, 2002.

S. Gavrilets. Perspective: models of speciation: what have we learned in 40 years? *Evolution*, 57(10):2197–2215, 2003.

S. Gavrilets. Models of speciation: where are we now? *Journal of heredity*, 105(S1): 743–755, 2014.

P.A. Gros, H. Le Nagard, and O. Tenaillon. The evolution of epistasis and its links with genetic robustness, complexity and drift in a phenotypic model of adaptation. *Genetics*, 182(1):277–293, 2009.

R. Guerrero, C. Muir, S. Josway, and L. Moyle. Pervasive antagonistic interactions among hybrid incompatibility loci. *PLoS Genetics*, 13(6):e1006817, 2017.

J. Haldane. The relative importance of principal and modifying genes in determining some human diseases. *Journal of Genetics*, 41:149–157, 1941.

J. Hey and C. Pinho. Population genetics and objectivity in species diagnosis. *Evolution*, 66(5):1413–1429, 2012.

D. Irwin. Assortative mating in hybrid zones is remarkably ineffective in promoting speciation. *The American Naturalist*, 195(6):E150–E167, 2020.

Y. Jiang, D. Bolnick, and M. Kirkpatrick. Assortative mating in animals. *The American Naturalist*, 181(6):E125–E138, 2013.

K. Kao, K. Schwartz, and G. Sherlock. A genome-wide analysis reveals no nuclear Dobzhansky-Muller pairs of determinants of speciation between S. cerevisiae and S. paradoxus, but suggests more complex incompatibilities. *PLoS genetics*, 6(7):e1001038, 2010.

J. Kitano and G. Okude. Causative genes of intrinsic hybrid incompatibility in animals and plants: what we have learned about speciation from the molecular perspective. *Evolutionary Journal of the Linnean Society*, 3(1):kzae022, 2024.

J. Kulmuni and A. Westram. Intrinsic incompatibilities evolving as a by-product of divergent ecological selection: Considering them in empirical studies on divergence with gene flow. *Molecular Ecology*, 26(12):3093–3103, 2017.

Y. Lee, C. Marsden, L. Norris, T. Collier, B. Main, A. Fofana, A. Cornel, and G. Lanzaro. Spatiotemporal dynamics of gene flow and hybrid fitness between the m and s forms of the malaria mosquito, anopheles gambiae. *Proceedings of the National Academy of Sciences*, 110(49):19854–19859, 2013.

L. Maisonneuve, C. Smadi, and V. Llaurens. Which cues are sexy? the evolution of mate preference in sympatric species reveals the contrasted effect of adaptation and reproductive interference. *Evolution Letters*, 8(2):283–294, 2024.

J. Mallet and S. Mullen. Reproductive isolation is a heuristic, not a measure: a commentary on westram et al., 2022. *Journal of Evolutionary Biology*, 35(9):1175–1182, 2022.

M. McGee, S. Borstein, J. Meier, D. Marques, S. Mwaiko, A. Taabu, M. Kishe, B. O'Meara, R. Bruggmann, L. Excoffier, et al. The ecological and genomic basis of explosive adaptive radiation. *Nature*, 586(7827):75–79, 2020.

M. McPeek and S. Gavrilets. The evolution of female mating preferences: differentiation from species with promiscuous males can promote speciation. *Evolution*, 60(10):1967–1980, 2006.

H. Muller. Isolating mechanisms, evolution and temperature. In *Biol. Symp*, volume 6, pages 71–125, 1942.

M. Nei. Modification of linkage intensity by natural selection. *Genetics*, 57(3):625, 1967.

P. Nosil. *Ecological speciation*. Oxford University Press, 2012.

P. Nouhaud, A. Blanckaert, C. Bank, and J. Kulmuni. Understanding admixture: haplodiploidy to the rescue. *Trends in Ecology & Evolution*, 35(1):34–42, 2020.

H. Orr. The population genetics of speciation: the evolution of hybrid incompatibilities. *Genetics*, 139(4):1805–1813, 1995.

B. Payseur and L. Rieseberg. A genomic perspective on hybridization and speciation. *Molecular ecology*, 25(11):2337–2360, 2016.

D. Powell, M. García-Olazábal, M. Keegan, P. Reilly, K. Du, A. Díaz-Loyo, S. Banerjee, D. Blakkan, D. Reich, P. Andolfatto, et al. Natural hybridization reveals incompatible alleles that cause melanoma in swordtail fish. *Science*, 368(6492):731–736, 2020.

D. Presgraves. A fine-scale genetic analysis of hybrid incompatibilities in *Drosophila*. *Genetics*, 163(3):955–972, 2003.

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2022. URL `https://www.R-project.org/`.

H. Rundle and P. Nosil. Ecological speciation. *Ecology letters*, 8(3):336–352, 2005.

M. Servedio and J. Hermisson. The evolution of partial reproductive isolation as an adaptive optimum. *Evolution*, 74(1):4–14, 2020.

J. Sobel, G. Chen, L. Watt, and D. Schemske. The biology of speciation. *Evolution*, 64 (2):295–315, 2010.

S. Stankowski and M. Ravinet. Defining the speciation continuum. *Evolution*, 75(6): 1256–1273, 2021.

R. Stelkens, K. Young, and O. Seehausen. The accumulation of reproductive incompatibilities in african cichlid fish. *Evolution: International Journal of Organic Evolution*, 64(3):617–633, 2010.

R. Stelkens, C. Schmid, and O. Seehausen. Hybrid breakdown in cichlid fish. *PloS one*, 10(5):e0127207, 2015.

T. Turner, M. Hahn, and S. Nuzhdin. Genomic islands of speciation in anopheles gambiae. *PLoS Biol*, 3(9):e285, 2005.

B. Wallace. Polymorphism, population size, and genetic load. *Population biology and evolution*, pages 87–108, 1968.

B. Wallace. Hard and soft selection revisited. *Evolution*, pages 465–473, 1975.

A. Westram, S. Stankowski, P. Surendranadh, and N. Barton. What is reproductive isolation? *Journal of evolutionary biology*, 35(9):1143–1164, 2022.

Hadley Wickham. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2016. ISBN 978-3-319-24277-4. URL `https://ggplot2.tidyverse.org`.

Wolfram Research, Inc. Mathematica v. 12. Champaign, Illinois, USA., 2019. https://www.wolfram.com.

R. Yamaguchi and Y. Iwasa. Reproductive character displacement by the evolution of female mate choice. *Evolutionary Ecology Research*, 15(1):25–41, 2013.