

RESEARCH ARTICLE

In search of the Goldilocks zone for hybrid speciation

Alexandre Blanckaert*, Claudia Bank

Instituto Gulbenkian de Ciência, Oeiras, Portugal

* blanckaert.a@gmail.com

Abstract

Hybridization has recently gained considerable interest both as a unique opportunity for observing speciation mechanisms and as a potential engine for speciation. The latter remains a controversial topic. It was recently hypothesized that the reciprocal sorting of genetic incompatibilities from parental species could result in hybrid speciation, when the hybrid population maintains a mixed combination of the parental incompatibilities that prevents further gene exchange with both parental populations. However, the specifics of the purging/sorting process of multiple incompatibilities have not been examined theoretically. We here investigate the allele-frequency dynamics of an isolated hybrid population that results from a single hybridization event. Using models of two or four loci, we investigate the fate of one or two genetic incompatibilities of the Dobzhansky-Muller type (DMIs). We study how various parameters affect both the sorting/purging of the DMIs and the probability of observing hybrid speciation by reciprocal sorting. We find that the probability of hybrid speciation is strongly dependent on the linkage architecture (i.e. the order and recombination rate between loci along chromosomes), the population size of the hybrid population, and the initial relative contributions of the parental populations to the hybrid population. We identify a Goldilocks zone for specific linkage architectures and intermediate recombination rates, in which hybrid speciation becomes highly probable. Whereas an equal contribution of parental populations to the hybrid population maximizes the hybrid speciation probability in the Goldilocks zone, other linkage architectures yield unintuitive asymmetric maxima. We provide an explanation for this pattern, and discuss our results both with respect to the best conditions for observing hybrid speciation in nature and their implications regarding patterns of introgression in hybrid zones.



OPEN ACCESS

Citation: Blanckaert A, Bank C (2018) In search of the Goldilocks zone for hybrid speciation. *PLoS Genet* 14(9): e1007613. <https://doi.org/10.1371/journal.pgen.1007613>

Editor: Jianzhi Zhang, University of Michigan, UNITED STATES

Received: May 18, 2018

Accepted: August 7, 2018

Published: September 7, 2018

Copyright: © 2018 Blanckaert, Bank. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The source code for the simulations can be found at <https://gitlab.com/evoldyn/four-loci>.

Funding: This research was supported by the Fundação Calouste Gulbenkian. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Author summary

Hybridization is observed ubiquitously in nature. Its outcome can range from extinction to the creation of new species. With respect to the latter, the probability of homoploid hybrid speciation, i.e. the formation of a new species as a result of a hybridization event without changes in the ploidy of the organism, is a hotly debated topic. Here, we analyze a minimal model for homoploid hybrid speciation, in which reproductive isolation is achieved by means of (postzygotic) Dobzhansky-Muller incompatibilities. When these

postzygotic genetic incompatibilities are resolved in the hybrid population, their reciprocal sorting can result in reproductive isolation from both parental populations, thus creating a hybrid species. We show that, in accordance with the current literature, hybrid speciation tends to be rare. However, specific arrangements of the genes responsible for reproductive isolation can make reciprocal sorting almost unavoidable and thus create barriers to the parental population in an almost deterministic matter. We discuss the implications of these results for hybrid speciation and patterns of introgression in nature.

Introduction

The role of hybridization in adaptation and speciation is an ongoing contentious question [1–10]. On the one hand, hybridization may serve as a source of genetic variation. Various examples of adaptive introgression have been reported (reviewed in [11]), and it has been argued that hybridization may provide the fuel for adaptive radiations [12]. On the other hand, gene flow between diverging populations may slow down or even reverse speciation either by purging isolating barriers or by one population swamping the other [13–15]. Thus, hybridization may act both as an engine of speciation and a boost to genetic variation, and as a detrimental mechanism that reduces population fitness and promotes extinction. This duality makes hybridization an important subject of study not only from an evolutionary but also a conservation biology point of view.

Hybrid speciation describes a scenario in which hybridization is essential for the formation of a “daughter” species that is isolated from both its parental species. The term “hybrid speciation” covers different scenarios that can be distinguished by the mechanism responsible for the buildup of reproductive isolation. In the case of polyploidization, individuals of the newly formed species possess more than one copy of each parents’ chromosomes. The parents can be of the same species (autopolyploidization, although the hybrid species tends to be outcompeted by the parental diploid [6]), or different ones (allopolyploidization), resulting in a single-step speciation event. In contrast, homoploid hybrid speciation (or recombinational speciation) corresponds to the formation of a hybrid species without a change in the ploidy level. For homoploid hybrid speciation to be possible, two criteria have to be fulfilled. On the one hand, there must be existing genetic isolating barriers between the parental species. On the other hand, these barriers have to be incomplete, such that sufficiently fit F1 hybrids can found the hybrid species. Despite this apparent paradox, numerous empirical cases of homoploid hybrid speciation have recently been reported [16–22]. Whether all of these represent true cases of homoploid hybrid speciation has been subject to debate. This debate has been led mainly around the definition of hybrid speciation and the resulting implications for the reported cases of empirical evidence [9, 10, 23]. However, to our knowledge there exists little work that has evaluated the probability of hybrid speciation theoretically.

Buerkle *et al.* [4] studied a specific case of hybrid speciation via two overlapping parental inversions. Their simulations suggested a rather narrow parameter range in which hybrid speciation is possible, and indicated that (among other restrictions) high fertility of F1 hybrids is necessary to produce a stable hybrid population, which, as a consequence, is only poorly isolated from its parental species. Moreover, Schumer *et al.* [24] studied the conditions for reciprocal sorting of genetic incompatibilities. A single genetic incompatibility can only isolate the hybrid population from one of its parental origins. However, if multiple DMIs exist between two species, in a hybrid population they might be resolved reciprocally with respect to the parental allelic origin, which can result in a hybrid species that is isolated from both parental

populations. Proposing this model, Schumer *et al.* [24] demonstrated via simulations that pairs of genetic incompatibilities can trigger hybrid speciation when there is a cost to both the ancestral genotype and the incompatibility-bearing genotypes. Based on a similar model, Comeault [25] recently investigated the impact of adaptive loci linked to genetic incompatibilities on the probability of homoploid hybrid speciation in a stepping-stone model.

Inspired by Schumer's model, we here provide a detailed analysis of the probability and dynamics of (reciprocal) sorting of classical Dobzhansky-Muller incompatibilities (DMIs [26–28]). A DMI consists of two (individually neutral or beneficial) alleles at different loci that are negatively epistatic, i.e., their combination is deleterious. The Dobzhansky-Muller model is arguably the most widely accepted model to explain the buildup of intrinsic postzygotic isolation in allopatric populations, and many examples of DMIs have been identified empirically [29–36]. Under this premise, such DMIs are expected to be the most prevalent type of genetic incompatibility that can be involved in reciprocal sorting and thus contribute to hybrid speciation. In our model, direct selection on the derived alleles is assumed to be absent or weak as compared with epistasis (in contrast to the fitness landscape in Schumer *et al.* [24]), and thus our model relies exclusively on classical DMIs [37]. We explain the differences between the two models in Supplementary Section S1. Note that we focus only on the genetic mechanisms responsible for hybrid speciation and ignore the potential role of ecological factors [38].

We identify several parameters that greatly influence the probability of hybrid speciation via DMIs. Specifically, we quantify how the population size, the initial contribution of parental alleles, and the linkage architecture affect the probability of hybrid speciation. As linkage architecture, we define the relative position of and recombination rate between the different loci involved in the hybrid incompatibilities that contribute to the species barriers (see also Fig 1). Consistent with Schumer *et al.* [24], we define hybrid speciation as the successful reciprocal sorting of incompatibilities, independent of the amount of isolation they confer. We discuss both weak and strong isolating barriers and consider recessive and codominant DMIs [39, 40], which differ considerably in their sorting patterns. Our results indicate that the linkage architecture of the DMIs plays an essential role, such that a specific architecture can make hybrid

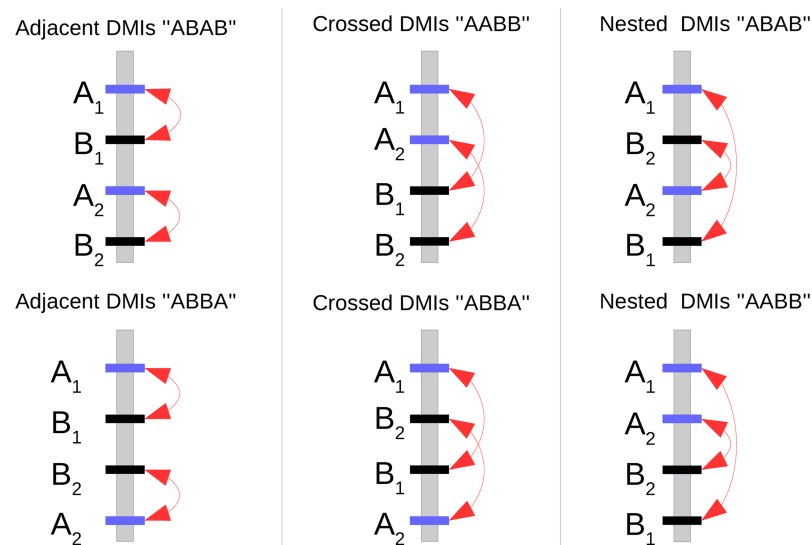


Fig 1. Illustration of all six different linkage architectures possible along a single chromosome. The A_k loci are given in blue and B_k in black. Red arrows show the incompatible interactions. The name of each linkage architecture is derived from the relative arrangement of the two incompatibilities and the order of the A and B loci.

<https://doi.org/10.1371/journal.pgen.1007613.g001>

speciation almost unavoidable, whereas exchanging only two loci may make hybrid speciation impossible for otherwise unaltered parameter values. Thus, we identify a Goldilocks zone of hybrid speciation, in which an interplay of population size, linkage architecture and the relative contribution of parental genomes can make hybrid speciation more likely than previously assumed.

Materials and methods

Simulations

Simulations are implemented in C++ (available at <https://gitlab.com/evoldyn/four-loci>) and each run ended when the population was monomorphic.

Results

Model

We study the evolution of a single population of constant size N in discrete generations. We model four diallelic loci, A_1, A_2, B_1, B_2 . At each locus, the derived allele is named after the locus, and the ancestral allele is named after its corresponding lower-case letter. Note that we do not detail here the two-locus model as it is fully included in the four-locus model. It can be obtained by keeping only loci A_1 and B_1 . Derived alleles at the different loci are under direct selection (soft selection), with α_k the (direct) fitness advantage of allele A_k over a_k , and β_k the fitness advantage of allele B_k over b_k . Selection happens in the diploid phase of the life cycle. In addition, negative epistasis, ϵ_k , which determines the strength of hybrid incompatibility, occurs in a pairwise fashion between the derived A_k and B_k alleles (with $k \in \{1, 2\}$). Dominance affects only the epistatic interactions, whereas the direct fitness effects are assumed to be multiplicative throughout the whole manuscript. We focus mainly on two cases of dominance of the epistatic interactions, which were proven representative of the general pattern [40]: a recessive scenario and a codominant scenario, illustrated in Fig 2. We introduce ϕ_k^n as a mathematical placeholder used to distinguish between the recessive and codominant scenario at the DMI k , with n the number of pairs of incompatible alleles in a genotype. Note that $n = 1, n = 2$ and $n = 4$ correspond to the H_0, H_1 and H_2 incompatibilities in Turelli & Orr [39]. Therefore, for a codominant DMI, $\phi_k^n = 1 \forall n \in \{0, 1, 2, 4\}$ while for a recessive DMI, the effect of epistasis is masked for the double heterozygote genotype, i.e. $\phi_k^1 = 0$ while $\forall n \in \{0, 2, 4\}, \phi_k^n = 1$.

The population is initially composed of two single genotypes, since it is assumed to result from a one-time secondary contact between two monomorphic parental populations 1 and 2; i_p denotes the contribution of the parental population 1 to the newly formed hybrid population. We assume that parental population 1 is fixed for the $A_1b_1A_2b_2/A_1b_1A_2b_2$ genotype and parental population 2 for $a_1B_1a_2B_2/a_1B_1a_2B_2$. The fitness of a genotype composed of haplotypes i and j is given by:

$$\omega^{ij} = \prod_{k=1}^2 (1 + \alpha_k)^{X_k^i + X_k^j} (1 + \beta_k)^{Y_k^i + Y_k^j} \left(1 + \phi_k^{(X_k^i + X_k^j)(Y_k^i + Y_k^j)} \cdot \epsilon_k \right)^{(X_k^i + X_k^j)(Y_k^i + Y_k^j)}, \quad (1)$$

where X_k^i is the number of alleles A_k in haplotype i and Y_k^i the number of alleles B_k in haplotype i . The full fitness landscape is illustrated in S1 Fig.

Mating is random. We assume that the parents generate an infinite pool of gametes from which zygotes are formed through multinomial sampling $M(2N, p_1, \dots, p_{16})$. Note that the deterministic case (i.e., in the absence of drift, $N \rightarrow \infty$) can be obtained by skipping the multinomial sampling step during zygote formation.

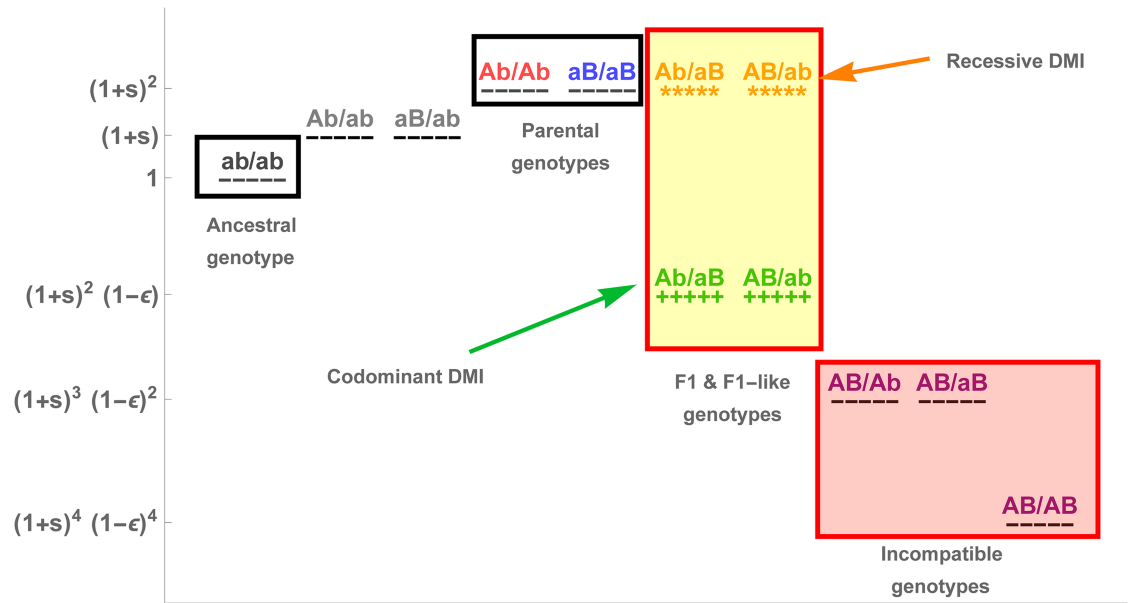


Fig 2. Fitness landscape of the 16 genotypes in the two-locus model, highlighting the effect of dominance of the incompatibility on the fitness of F1 hybrids. For simplicity, we illustrate the case of $\alpha = \beta = s$, where s is the selective advantage of a derived allele in the ancestral background. Note that there are only 10 genotypes represented here, as we do not distinguish the parental origins of each haplotype.

<https://doi.org/10.1371/journal.pgen.1007613.g002>

As introduced above, hybrid speciation is defined as the fixation of a haplotype that is incompatible with both parental haplotypes, see Table 1. Indeed, if an individual homozygous for the $A_1b_1a_2B_2$ haplotype is backcrossed with an individual from, e.g., parental population 1, then the second DMI is expressed either in the F1 generation (codominant case) or in the F2 generation (recessive case). Similarly, introduction of an $A_1b_1a_2B_2/A_1b_1a_2B_2$ individual in the parental population 2 leads to the expression of the first DMI. This definition corresponds to an early stage speciation mechanism, leading to a hybrid population that is only partially isolated from both parental populations. Note that full isolation is impossible in this setting, as barriers responsible for full reproductive isolation would also prevent the formation of the hybrid population in the first place.

We consider all possible linkage architectures formed by the two DMIs; they are illustrated in Fig 1. There are 6 different ways to arrange the 4 loci along a single chromosome (assuming the chromosome does not have an orientation). The two DMIs can be “Adjacent”, “Crossed”,

Table 1. Classification of possible haplotypes for the “Adjacent ABAB” linkage architecture.

| | |
|--|--|
| Ancestral haplotype | $a_1b_1a_2b_2$ |
| Parental pop. 1 haplotype | $A_1b_1A_2b_2$ |
| Parental pop. 2 haplotype | $a_1B_1a_2B_2$ |
| Hybrid haplotypes | $A_1b_1a_2B_2$ or $a_1B_1A_2b_2$ |
| “Epistasis-free” F2 haplotype | $A_1b_1a_2b_2$ or $a_1B_1a_2b_2$ or $a_1b_1A_2b_2$ or $a_1b_1a_2B_2$ |
| 1 st incompatibility haplotypes | $A_1B_1a_2b_2$ or $A_1B_1A_2b_2$ or $A_1B_1a_2B_2$ |
| 2 nd incompatibility haplotypes | $a_1b_1A_2B_2$ or $a_1B_1A_2B_2$ or $A_1b_1A_2B_2$ |
| Double-incompatibilities haplotype | $A_1B_1A_2B_2$ |

<https://doi.org/10.1371/journal.pgen.1007613.t001>

or “Nested” (Fig 1). The recombination rate between adjacent loci X and Y is given by $0 \leq r_{XY} \leq 0.5$. The recombination rate between non-adjacent loci X and Y, separated by a single locus W, is given by $r_{XY} = r_{XW}(1 - r_{WY}) + r_{WY}(1 - r_{XW})$. If the four loci are spread across multiple chromosomes, this represents a special case of the single-chromosome scenario presented above, in which one or more recombination rates are set to 0.5. If not otherwise specified, we assume that all loci are located on different chromosomes, i.e. $r_{XY} = 0.5$.

Resolution of a single DMI

In the first part, we focus on the resolution of a single DMI following the formation of the hybrid population. We define a DMI as “resolved” when either allele *A* or *B* is lost, such that there is no genetic conflict left in the population. With a single DMI, hybrid speciation according to our definition is impossible, because one of the negatively interacting partners *A* and *B* in the DMI will invariably be lost (see also below). Thus, maintaining a genetic barrier to both parental species is impossible. Nevertheless, the study of a single incompatibility pair is necessary to understand which properties can be extrapolated to multiple incompatibilities. We characterize the resolution of the genetic conflict resulting from the contact between two diverged populations by quantifying the probability of fixation of the different haplotypes, the time of resolution of the DMI (i.e., the time until at least one of the incompatible alleles is lost) and the time to fixation of a single haplotype. In this section, we only focus on the loci A_1 and B_1 and drop the indices as they do not carry any information.

Dynamics following secondary contact. In a single randomly mating population such as the hybrid population we consider here, a DMI cannot be maintained unless directional selection is strong as compared with the epistatic effect of the incompatibility [40]. This is because the formation of hybrid individuals initially leads to selection against both derived haplotypes. These haplotypes suffer from the incompatibility, either directly by forming an unfit hybrid genotype or indirectly through the production of unfit offspring. In contrast, the ancestral haplotype has an advantage as soon as it appears, and rises in frequency, because it only forms compatible genotypes and produces compatible offspring (if the proportion of incompatible *AB* haplotypes in the population remains low). As soon as the ancestral haplotype becomes frequent or either of the derived haplotypes becomes rare, this marginal advantage disappears, and the ancestral type will either be swamped by the more frequent derived type (in the case of direct selection acting on the derived alleles, i.e., if $\alpha, \beta > 0$), or segregate neutrally (if $\alpha, \beta = 0$). The incompatibility is usually resolved in favor of the more frequent derived allele (if they have similar fitness). Thus, one main determining factor is the initial frequency ratio between the two derived alleles (Supplementary Section S2.1). Both direct selection and codominance of the incompatibility reduce the impact of genetic drift (i.e., the outcome converges to the deterministic case). Indeed, once the DMI is resolved, selection increases the probability of fixation of a single derived allele [41, 42]. The codominance of the incompatibility shortens the time required to resolve the DMI (Fig 3), and therefore reduces the time spent at low frequencies, where loss of the derived alleles because of drift is a likely outcome (see Supplementary Section S2.2).

Recombination has opposite effects under different dominance schemes. Recombination has a dual impact on the outcome of a hybridization event, depending on the dominance of the DMI, as illustrated in Fig 3 for haplotype *Ab*. Recombination breaks the association between the alleles of the parental haplotype and therefore leads to the formation of both the incompatible haplotype *AB* and the ancestral haplotype *ab*. On the one hand, this allows the expression of the incompatibility through the formation of the *AB* haplotype, which leads to faster sorting of the derived alleles. On the other hand, building a genotype with the ancestral

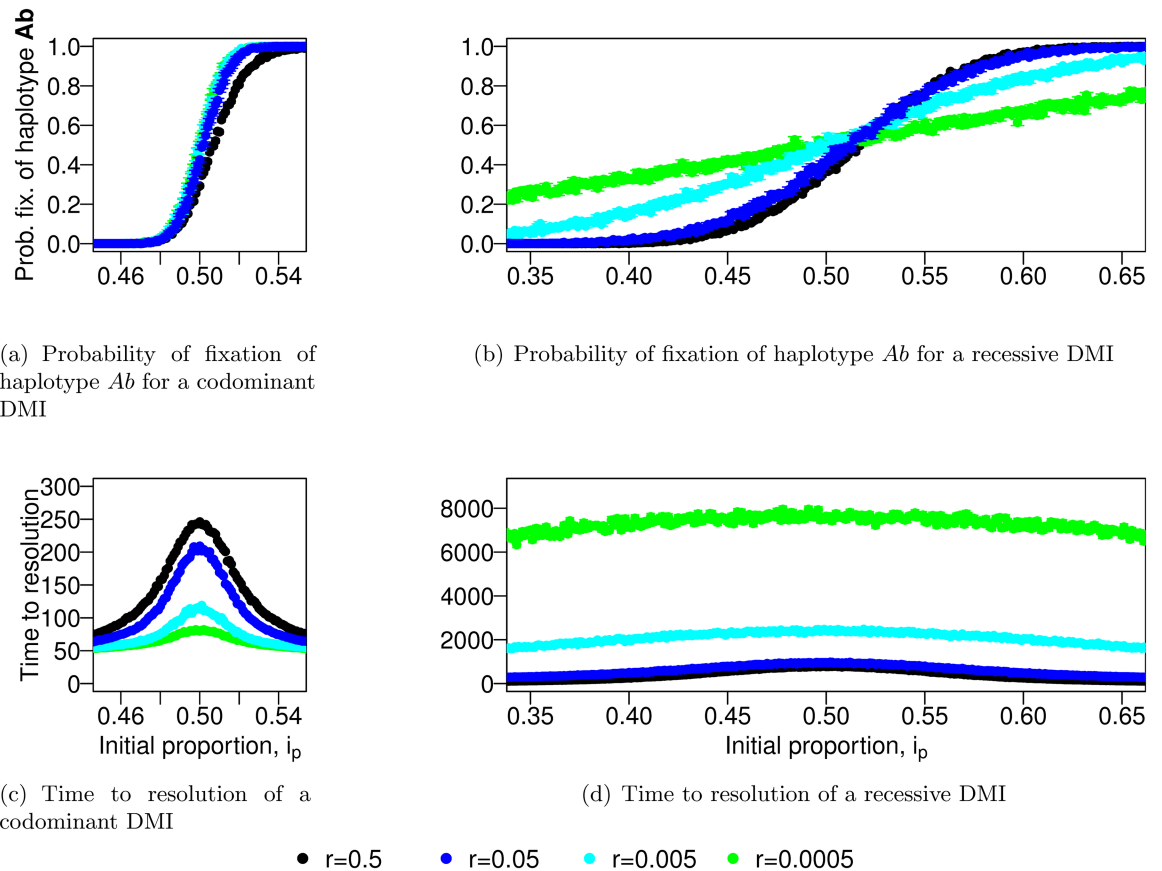


Fig 3. Recombination slows down the resolution of a codominant DMI whereas it speeds DMI resolution up for a recessive one. We represent the probability of fixation of the *Ab* haplotype (panels (a,b)) for different recombination rates and different dominance schemes (codominant (a,c), recessive (b,d)). In panels (c,d), we illustrate the time to resolve the genetic conflict (i.e., either allele *A* or *B* is lost). Each value is obtained from 1000 independent simulations. Note the much larger y-axis scale (x30) in panel (d). Parameters used: $\alpha = \beta = 0.001$, $\epsilon = -0.2$, $N = 5000$.

<https://doi.org/10.1371/journal.pgen.1007613.g003>

haplotype protects both parental haplotypes from suffering from the genetic incompatibility, which leads to slower sorting of the derived alleles. The balance between these two effects is different between a recessive and a codominant DMI, which results in an opposite behavior, as illustrated in the following.

In the recessive case, recombination is necessary for the expression of the incompatibility, because only production of *AB* haplotypes unmasks the epistatic interaction. The faster these haplotypes are produced, the stronger (epistatic) selection acts, which leads to a faster resolution of the DMI. An increase in recombination therefore always accelerates the resolution of a recessive DMI. This reduces the time the derived alleles spend at low frequency, which makes them less susceptible to being lost through genetic drift. This, in turn, reduces the probability that the ancestral haplotype becomes fixed.

In the codominant case, the incompatibility is already expressed in the F1 generation. Recombination is not necessary to express the incompatibility and therefore slows down the resolution of the DMI, as the ancestral haplotype prevents the effective purging of the parental haplotypes through the formation of *ab/Ab* or *ab/AB* individuals. In this situation, both derived alleles are driven to lower frequencies than in the recessive model, which makes them more likely to be both lost through genetic drift, resulting in the fixation of the ancestral haplotype.

Resolution of two DMIs and hybrid speciation

Expanding from what we learned above, we now consider what happens when two incompatibilities exist between the parental populations. In contrast to the case of a single DMI, a new evolutionary outcome, namely hybrid speciation, becomes feasible with more than one DMI. As “hybrid speciation”, we denote the reciprocal sorting of the two DMIs, i.e. fixation of either alleles A_1 and B_2 or A_2 and B_1 . Such a hybrid population will then carry isolating barriers to both parental populations.

Isolation of the hybrid population by reciprocal sorting of two DMIs. Given the observed shape of the fixation probability of a derived allele in the case of a single DMI as a function of the initial contribution of both parental populations (Fig 3), hybrid speciation should be observable only around symmetric contact, and this condition should be more stringent for codominant incompatibilities than recessive ones. In Fig 4, we test this expectation by comparing the probability of hybrid speciation for two DMIs that are located on separate chromosomes (with A_1 and B_1 on the first chromosome and A_2 and B_2 on the second one; colored dots in Fig 4) with the expected probability of resolving two independent single DMIs for opposite derived alleles (e.g. first DMI resolved towards allele A and the second one for allele B ; black dots). In the recessive case, the prediction for independent DMIs matches the hybrid speciation probability. In the codominant case, the independent expectation overestimates the probability of hybrid speciation. This can be explained by the faster resolution of the DMIs in the codominant model, which, even in the case of free recombination, leaves insufficient time for the two DMIs to become uncoupled and independently resolved, as the A_1 and A_2 loci start in maximum linkage disequilibrium. In the codominant case, this effect is amplified at low recombination rates as, in that case, the resolution of the DMIs happens even faster (S3 Fig), therefore preserving more of the initial linkage disequilibrium. This leads to a positive correlation between the fixation of the different A_i alleles (as well as B_j alleles), S2 Fig. In the recessive case, the resolution of the two DMIs remains independent as it takes much longer to resolve any single recessive DMI.

S3 Fig illustrates the mean time to resolve both DMIs in opposite directions conditioned on the outcome being hybrid speciation. Recombination has the same effect on the resolution of two DMIs as it did for a single one: it accelerates the resolution of recessive DMIs and slows down the resolution of codominant DMIs. However, the average resolution time is not affected by the initial proportion of the parental species; only trajectories that resolve slowly, which ensure that the initial linkage disequilibrium has been broken, can contribute to hybrid speciation, and we are conditioning on this outcome.

Due to the complexity of the dynamics, the possibility to obtain analytical results was very limited. However, we were able to obtain a condition for the (deterministic) local stability of the hybrid species with respect to back mutation or single migration events (see Supplementary Section S4). Most notably, we find that the condition for local stability of the monomorphic equilibrium that contains the reciprocally sorted haplotypes is independent of the dominance of the incompatibilities.

The linkage architecture determines which alleles survive. Fig 5 illustrates the effect of recombination and the linkage architecture on hybrid speciation, when all loci are on the same chromosome (as opposed to one DMI per chromosome, as illustrated in Fig 4). As mentioned above, for codominant DMIs, recombination, on the one hand, allows the formation of the hybrid haplotype and helps to reduce the initial linkage disequilibrium. On the other hand, it slows down the resolution of the DMI through the formation of compatible haplotypes. Depending on the balance between these two effects, recombination impacts the probability of hybrid speciation differently.

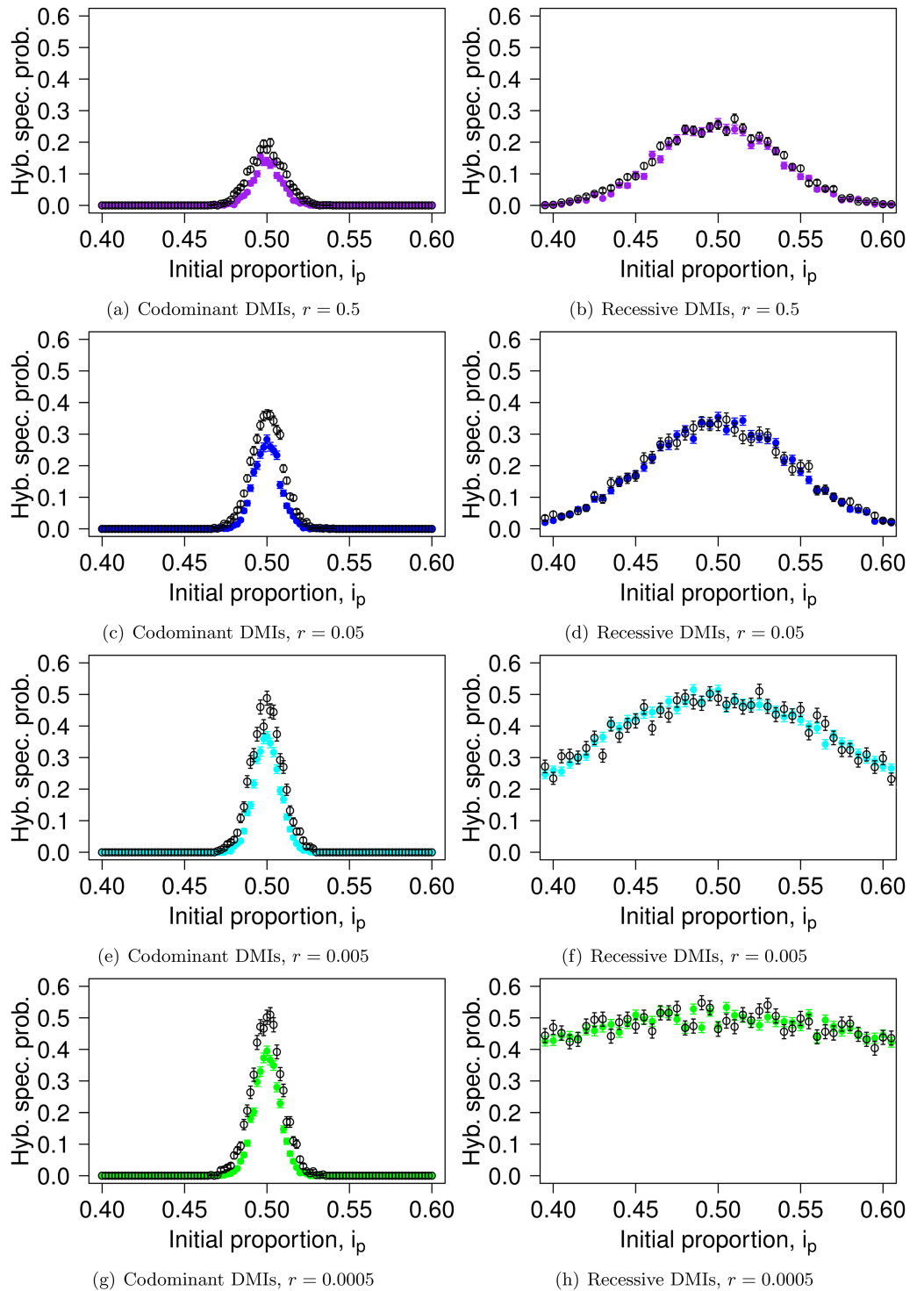


Fig 4. Hybrid speciation probability for codominant (panels (a,c,e,g)) and recessive (panels (b,d,f,h)) DMIs. The colored dots correspond to the probability of hybrid speciation for two DMIs situated on different chromosomes ($r_{23} = 0.5$). The recombination rate between the interacting loci is indicated below each panel ($r_{12} = r_{34} = r$): panels (a,b), $r = 0.5$; panels (c,d), $r = 0.05$; panels (e,f), $r = 0.005$; panels (g,h), $r = 0.0005$. The black dots correspond to the predicted hybrid speciation probability based on the resolution of a single DMI. The fast resolution of the codominant DMIs leads to a correlation between their fate, which makes hybrid speciation less likely than the independent expectation predicts. Parameters used are $\alpha_i = \beta_i = 0.001$, $N = 5000$, $\epsilon = -0.2$. Each dot is obtained from 1000 replicates.

<https://doi.org/10.1371/journal.pgen.1007613.g004>

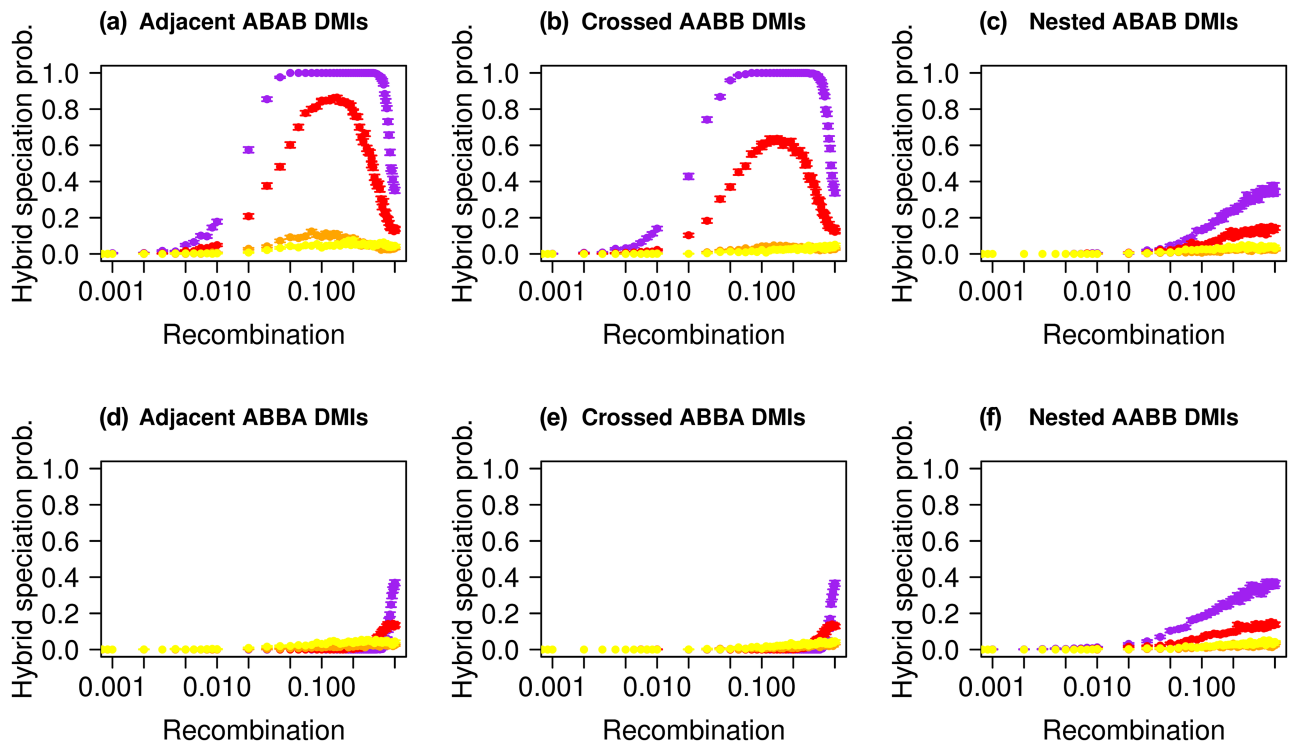


Fig 5. Hybrid speciation probability is a nonlinear function of recombination. We consider that all four loci have the same selective advantage ($\alpha_k = \beta_j = 0.001$) and are equidistant along a single chromosome. The hybrid speciation probability is plotted for different population sizes: yellow corresponds to $N = 50$, orange to $N = 500$, red to $N = 5000$ and purple to $N = 50000$. Epistasis ($\epsilon = -0.2$) is here codominant but we obtain qualitatively similar results for recessive incompatibilities, see S10 Fig. The contribution of both parental populations is symmetric ($i_p = 0.5$). Each panel corresponds to a different linkage architecture: (a) “Adjacent ABAB”; (b) “Crossed AABB”; (c) “Nested ABAB”; (d) “Adjacent ABBA”; (e) “Crossed ABBA”; (f) “Nested AABB”.

<https://doi.org/10.1371/journal.pgen.1007613.g005>

Assuming a symmetric contact, we observe that two of the linkage architectures, the “Adjacent ABAB” and “Crossed ABAB” architectures exhibit a non-monotonic behavior with maximum hybrid speciation probability for intermediate recombination rates. These linkage architectures have in common that the loci A_1 and B_2 are located at the ends of the chromosomal region that contains the four focal loci. Hybrid speciation becomes almost certain for large population sizes and indeed corresponds to the deterministic outcome (i.e. in the absence of drift) for these two linkage architectures. More precisely, we observe a local maximum of the hybrid speciation probability for recombination rates around $r = 0.1$. The “Adjacent ABAB” and “Crossed ABAB” architectures, which show this behavior, are characterized by a higher marginal fitness of the A_2 and B_1 alleles compared to the other alleles in the deterministic case, which promotes hybrid speciation. For all other architectures either A_1 and A_2 or A_2 and B_2 have the highest marginal fitness. The higher marginal fitness stems from the production of the $a_1B_1a_2b_2$ and $a_1b_1A_2b_2$ haplotypes in the F2 generation (for the “Adjacent ABAB” architecture) that are relatively free of epistasis. The outcomes of a single recombination event per genome for all 6 architectures are given in Fig 6 and illustrate how the “Adjacent ABAB” and “Crossed ABAB” architectures stand out in the production of the haplotypes that are needed for hybrid speciation. Importantly, recombination is necessary to generate these haplotypes, but too much recombination will cancel their advantage. Indeed, for $r = 0.5$, all haplotypes are produced at the same frequency in the absence of selection. The dual effect of recombination leads therefore to the observed maximum in the hybrid speciation probability

| Genetic architecture | F1 hybrid | Single recombination event between: | | |
|----------------------|-----------|-------------------------------------|----------------------------------|----------------------------------|
| | | loci 1 and 2 | loci 2 and 3 | loci 3 and 4 |
| Adjacent ABAB | | $A_1B_1a_2B_2$ $a_1b_1A_2b_2$ | $A_1b_1a_2B_2$ $a_1B_1A_2b_2$ | $A_1b_1A_2B_2$ $a_1B_1a_2b_2$ |
| Adjacent ABBA | | $A_1B_1B_2a_2$ $a_1b_1b_2A_2$ | $A_1B_1b_2a_2$ $a_1b_1B_2A_2$ | $A_1b_1b_2a_2$ $a_1B_1B_2A_2$ |
| Crossed AABB | | $A_1a_2B_1B_2$ $a_1A_2b_1b_2$ | $A_1A_2B_1B_2$ $a_1a_2b_1b_2$ | $A_1A_2b_1B_2$ $a_1a_2B_1b_2$ |
| Crossed ABBA | | $A_1B_2B_1a_2$ $a_1b_2b_1A_2$ | $A_1B_2b_1a_2$ $a_1b_2B_1A_2$ | $A_1b_2b_1a_2$ $a_1B_2B_1A_2$ |
| Nested AABB | | $A_1a_2B_2B_1$ $a_1A_2b_2b_1$ | $A_1A_2B_2B_1$ $a_1a_2b_2b_1$ | $A_1A_2b_2B_1$ $a_1a_2B_2b_1$ |
| Nested ABAB | | $A_1B_2a_2B_1$ $a_1b_2A_2b_1$ | $A_1b_2a_2B_1$ $a_1B_2A_2b_1$ | $A_1b_2A_2B_1$ $a_1B_2a_2b_1$ |

Fig 6. Haplotypes produced in the F2 breakdown, assuming a single recombination event, explain how different linkage architectures leads to different outcomes for the same loci. By identifying the relatively “epistasis-free” haplotype formed, one can infer whether hybrid speciation may be a likely outcome. In blue, we highlight these “epistasis-free” haplotypes that are important for hybrid speciation and in yellow those that are important for fixation of the parental haplotype from population 1.

<https://doi.org/10.1371/journal.pgen.1007613.g006>

for intermediate recombination rates. When the DMIs are located on two different chromosomes (as in Fig 4), this effect does not appear. Indeed, while recombination still breaks linkage disequilibrium, it no longer generates the relatively “epistasis-free” haplotype and therefore leads to a monotonous increase in the hybrid speciation probability with increasing recombination rate. This behavior, specific to the “Adjacent ABAB” and “Crossed AABB” linkage architectures, is observed both for codominant and recessive DMIs.

As illustrated in S10 Fig, the recessive case is qualitatively similar to the codominant one. We recover the distinctive pattern between linkage architectures, where the “Adjacent ABAB” and “Crossed AABB” architectures are more likely to generate hybrid speciation for intermediate recombination rates. However, for the “Adjacent ABBA” and “Crossed ABBA” linkage architectures, the recessive case differs from the codominant by the existence of two local maxima for the hybrid speciation probability as a function of recombination. These two architectures are characterized by an indirect selective advantage of one of the two parental haplotypes over the other, as the partially derived haplotypes $A_1b_1b_2a_2$ and $a_1b_1b_2A_2$ are more likely to be formed than their counterparts ($a_1B_1b_2a_2$ or $a_1b_1B_2a_2$, see Fig 6), which

leads to a slightly higher marginal fitness of the $A_1b_1b_2A_2$ haplotype compared to $a_1B_1B_2a_2$. The first maximum is obtained at large intermediate recombination rates; it corresponds to the one observed for codominant DMIs. However, a second one can be observed at lower recombination rate if the population size is large enough. It results from a subtle balance between drift, recombination and selection, which we explain in detail in Supplementary Section S3.1.

The hybrid speciation probability for codominant versus recessive DMIs differs significantly when considering lethal incompatibilities (see Supplementary Section S3.1). Hybrid speciation becomes impossible for codominant DMIs because no viable hybrids can be produced. This is not the case for recessive incompatibilities, as they can partially escape the strong selection against hybrids. In fact, due to the masking effect provided in F1 and F1-like genotypes, we observe an almost indistinguishable pattern for deleterious ($\epsilon = -0.2$) and lethal ($\epsilon = -0.99$) recessive DMIs, see S14 Fig. Similarly, the time to hybrid speciation is similar between the deleterious and lethal recessive cases, see Supplementary Section S3.3.

Fig 5 also illustrates the impact of the population size on the outcome. In general, a larger population size results in a higher probability of hybrid speciation. This is especially true when the deterministic outcome corresponds to hybrid speciation (i.e., for the “Adjacent ABAB” and “Crossed AABB” architectures). In large populations, derived alleles are less likely to be lost during the reciprocal sorting of the genetic incompatibilities. The main exception to this rule exists when the deterministic outcome is the fixation of one parental haplotype. In that case, an intermediate population size will maximize the likelihood of hybrid speciation, as illustrated in Fig 5 for the “Adjacent ABBA” and “Crossed ABBA” architectures (see also S10 Fig). This intermediate value corresponds to a balance between a strong drift regime in which the ancestral and “epistasis free” haplotypes are most likely to fix, and the deterministic regime in which the $A_1b_1b_2A_2$ parental haplotype fixes.

Symmetric contact is not always the best condition for hybrid speciation. Fig 5 was obtained for $i_p = 0.5$, i.e. when both parental populations contribute equally to the hybrid population. It corresponds to the case that is the most frequently investigated in models. Fig 7 illustrates what happens when we release this assumption. From the single-DMI dynamics, one would expect a decrease in the hybrid speciation probability as illustrated in Fig 5. This is not always true. Depending on the linkage architecture, the probability of hybrid speciation may be higher for asymmetric contributions from the parental populations. This phenomenon is also observed for intermediate recombination rates; thus, only a consideration of dominance scheme, recombination rate, and symmetry together allows for an accurate statement on the hybrid speciation probability (see Fig 7). Fig 6 provides us with an explanation for the observed pattern: for intermediate recombination rate ($r \approx 1/3$), there is on average one recombination event per haplotype per generation. For the two architectures concerned (“Adjacent ABBA” and “Crossed ABBA”), in this scenario and with perfect symmetry, both alleles A_1 and A_2 have a marginal fitness that is slightly higher than alleles B_1 and B_2 (S4 Fig), which leads to the fixation of the parental haplotype $A_1b_1b_2A_2$ in the deterministic case. Therefore, a lower initial frequency of these alleles at the initial contact balances this selective advantage, which results in larger hybrid speciation probabilities than under symmetry. This behavior is only observed for the two architectures discussed above (“Adjacent ABBA” and “Crossed ABBA”). For the other architectures, the two derived alleles that have a slight indirect selective advantage are A_2 and B_1 for “Adjacent ABAB” and “Crossed AABB” (which correspond to the cases of high probabilities of hybrid speciation) or A_2 and B_2 for the two “Nested” architectures. In both cases, since the symmetry between the A and B alleles is respected, hybrid speciation is most likely at $i_p = 0.5$.

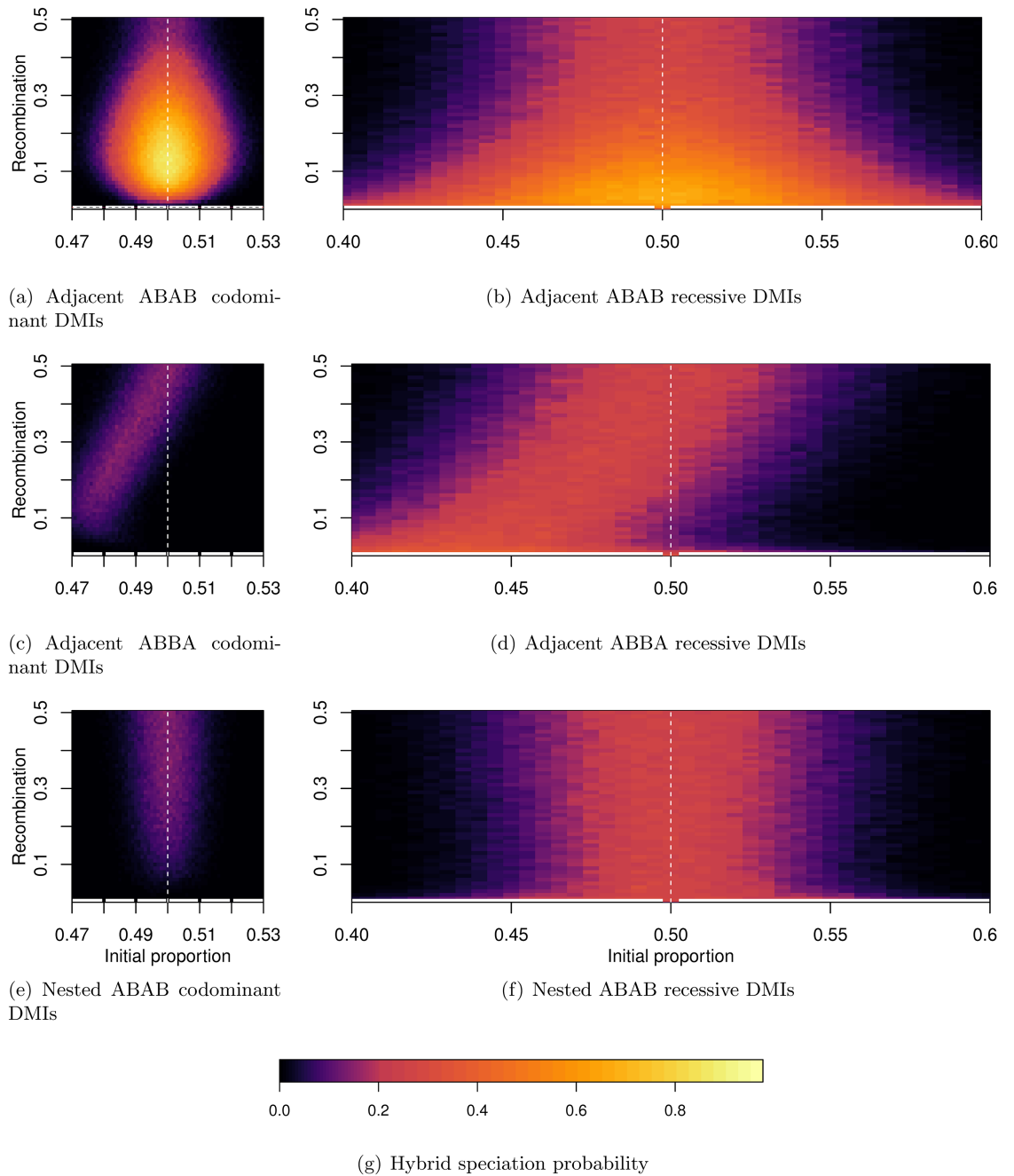


Fig 7. Probability of hybrid speciation for both recessive and codominant DMIs as a function of the recombination rate between the different loci (they are all equidistant) and the initial contribution of both parental species. Different linkage architectures generate an unexpected pattern: for the “Adjacent ABAB” architecture, we observe a Goldilocks zone for hybrid speciation; for the “Adjacent ABBA”, the probability of hybrid speciation is no longer symmetric along the $i_p = 0.5$ axis (white dashed line). Each panel corresponds to a given linkage architecture and a dominance scheme: (a) “Adjacent ABAB”, codominant DMIs; (b) “Adjacent ABAB”, recessive DMIs; (c) “Adjacent ABBA”, codominant DMIs; (d) “Adjacent ABBA” recessive DMIs; (e) “Nested ABAB”, codominant DMIs; (f) “Nested ABAB” recessive DMIs.

<https://doi.org/10.1371/journal.pgen.1007613.g007>

Discussion

We here characterized the purging process of single and multiple DMIs upon formation of an isolated hybrid population. Specifically, we quantified the effects of the linkage architecture and the dominance of the epistatic interactions on the reciprocal sorting of incompatibilities, which has been proposed as a mechanism to induce homoploid hybrid speciation. We found that for the exact same set of loci, their order along the chromosome can increase the probability of observing hybrid speciation from unlikely to almost certain. We demonstrate that the main determinant of this pattern is which haplotypes are formed during the F2 breakdown. For the linkage architectures that promote hybrid speciation, there exists a Goldilocks zone in which an intermediate recombination rate maximizes the hybrid speciation probability. In addition, we show that symmetric contact of incompatible loci that are under equal selection pressure does not always generate the highest probability of hybrid speciation, and that this result cannot be predicted from the study of independent DMIs. Finally, hybrid speciation with lethal codominant DMIs is impossible, whereas for recessive DMIs, in which the F1 generation does not suffer a fitness disadvantage, reciprocal sorting is similarly probable with intermediate and strong epistasis.

Linkage architecture as major determinant of hybrid speciation probability

Abbott *et al.* [7] recently emphasized that “an important challenge in studies of hybrid speciation is to ask whether there is an ‘optimal’ genetic distance for homoploid hybrid speciation [3, 43].” Although Abbott *et al.* [7] were arguably referring to the degree of divergence and thus, to the number and strength of isolating barriers that have established between two species, our study adds an additional important factor to their list: the linkage architecture of the isolating barriers, and the recombination rate between them. Specifically, our results demonstrate that intermediate recombination rates and specific linkage architectures maximize the probability of hybrid speciation.

We can speculate whether the presence of multiple DMIs should increase the probability of hybrid speciation. Based on our results, we believe that it should depend on the nature of the incompatibilities: additional recessive DMIs should increase the hybrid speciation probability while more codominant DMIs should reduce it. Firstly, we showed that for codominant DMIs, the fixation of derived alleles of the same parental population is correlated (S2 Fig), which hinders their reciprocal sorting even if they are located on different chromosomes. Hence, recombination is not sufficient to decrease the initially existing linkage disequilibrium. Adding additional codominant incompatibility loci will result in stronger selection against F1 individuals which strengthens the correlation of parental alleles, and reduces the probability of reciprocal sorting. Secondly, we have shown that for lethal codominant DMIs (S14 Fig), hybrid speciation is impossible. Extrapolating from these two observations, we propose that this effect will outpace the increase in hybrid speciation probability due to having more chances to have at least one pair of reciprocal sorting.

On the other hand, in the recessive model, F1 hybrids do not suffer a fitness cost, and the fixation of derived alleles from the same parental origin is uncorrelated. In addition, stronger epistasis does not affect the probability of hybrid speciation in the recessive model (S10 and S14 Figs). Thus, having more than two recessive DMIs should increase the chances that at least two are reciprocally sorted, which is sufficient for hybrid speciation according to our definition.

Our results can be discussed in the context of parapatric speciation and the role of genomic islands of divergence [44, 45]. According to the respective theory [46], during the speciation process and in presence of gene flow, islands of divergence are formed around the first genes

involved in reproductive isolation. These genes will reduce gene flow locally around them, which favors the accumulation of weakly locally adapted mutations in their vicinity, as well as incompatible genes, reinforcing and extending islands of divergence. For hybrid speciation according to our model, the existence of such islands implies that many derived alleles A_k may be found together on the same island, as locally adaptive loci tend to be rearranged in clusters [47]. In such cases the linkage disequilibrium between the different A_k alleles will be harder to break, which makes reciprocal sorting less probable. From the perspective of this model, hybrid speciation should therefore be most likely early during speciation, when no strong islands of divergence have formed yet. We further note that our model considers a single hybridization event without any further gene exchange with both parental populations, which resembles the colonization of a new environment. Continuous gene flow between populations, which is often a key feature of studies of genomic islands of speciation, should further reduce the probability of hybrid speciation [25], because migration creates selective pressure against the hybrid haplotypes.

The probability of hybrid speciation and reciprocal sorting in nature

Our results imply that although specific linkage architectures may indeed induce hybrid speciation with high probability, it remains on average unlikely. Other factors that we neglect in the minimal model presented here (such as polymorphic or complex DMIs, and ecological factors) may alter the hybrid speciation probability. However, our results are consistent with the scarcity of putative cases of homoploid hybrid speciation observed in nature [16–22]. Recently, Runemark *et al.* [48] reported that the Italian sparrow hybrid species resulted from multiple occurrences of hybridization between the Spanish and House sparrow along the Mediterranean Sea, which is in concordance with our observation of a Goldilocks zone. The observed resolution of all hybridization events towards a single mitochondrial origin (i.e. all Italian sparrows possess the mitochondrial DNA of House sparrows) indicates that either the mitochondria played an important role with respect to the sorting of the incompatibilities, or that there is an asymmetry in the formation of the different hybrid populations, in which Spanish sparrow males mate with House sparrow females.

The nature of genetic incompatibilities

Both theoretical considerations and empirical evidence suggest that most DMIs should be recessive [29, 49]. However, any kind of dominance pattern of the epistatic interactions can in theory exist [50]. Here, we showed that single codominant DMIs are resolved much faster than recessive ones. Therefore, when non-equilibrium populations are sampled, the excess of recessive incompatibilities may not necessarily reflect the true proportion of recessive incompatibilities but rather a sampling bias.

Genetic isolation of the hybrid population

Homoploid hybrid speciation relies on an apparent paradox: it requires that the existing genetic barriers are strong enough for the hybrid population to be isolated from both parental populations, while the same barrier needs to be weak enough to allow the formation of the hybrid population in the first place. We find that the reciprocal sorting of the two DMIs leads to a significant barrier for codominant DMIs (see S5 Fig) regardless of the genetic and linkage architecture of the incompatible loci. On the contrary, recessive DMIs only reduce the introgression probability significantly when the incompatibilities are lethal and the incompatible loci are far apart from each other. The measure for reproductive isolation we used here, the probability of introgression of an unlinked neutral allele, is conservative compared to measures

at the barrier loci themselves [40, 51, 52]. However, S5 Fig illustrates that early isolation of a hybrid population from its parental sources is possible for both dominance schemes. Nevertheless, in contrast to the polyploidisation scenario, complete reproductive isolation by means of homoploid hybrid speciation cannot be an immediate outcome but may happen by building on the initial reproductive isolation that reciprocal sorting confers.

The time to hybrid speciation

We specifically evaluated the timing of two events during the process of hybrid speciation. The first is when the genetic conflict caused by the DMIs is resolved by losing one of the interacting partners in each incompatibility pair; we call this the resolution time. From this point onwards, there are no epistatic interactions left in the population and the evolution of the remaining genotypes is dictated only by genetic drift and direct selection. The second event is the time at which all polymorphism is lost; conditioned on reciprocal sorting we call this the time to hybrid speciation. Only then, the isolating barrier to the parental populations is fixed in the population.

We find that initially, codominant DMIs are resolved faster, which leads to a short resolution time. This process is even faster with low recombination, because the lack of recombination decreases the probability to break the epistatically interacting haplotypes. (This also leads to a low probability of hybrid speciation for codominant DMIs with low recombination.) However, after resolution of the DMI, the derived alleles still need to become fixed in order for hybrid speciation to occur. This process is usually faster for recessive DMIs and compensates for the slower initial sorting, such that both codominant and recessive DMIs lead to similar times to hybrid speciation (see S15 and S16 Figs).

Schumer *et al.* [24] emphasized that hybrid speciation can sometimes happen quickly. Somewhat contradictory, we find that the time to hybrid speciation (scaled by population size) increases with smaller population sizes, and that there is relatively little variation in the times to hybrid speciation for constant population size (S15 and S16 Figs). For example, for a population of size $N = 50$, the time to hybrid speciation is around $5N = 250$ generations (i.e., longer than the average fixation time of a neutral allele) independent of the linkage architecture of the DMI. For a population of size $N = 50000$, it is only around $0.1N = 5000$ generations. That is because both epistatic and direct selection act more efficiently in large populations.

Population size and selection

In our model, we consider populations of constant size. Relaxing this assumption (i.e. switching from soft selection to hard selection), one would expect hybrid speciation to be less likely for at least two reasons. First, selection against the different derived alleles in the early purging phase is stronger; indeed, with soft selection the effect of a mutation is weighted by the mean fitness of the population. Therefore, in maladapted populations, the effect of deleterious mutations is slightly dampened. Second, the expected decrease in population size that is associated with the purging phase increases the impact of drift, which means that reciprocal sorting is less likely even with favorable linkage architectures (as selection is not strong enough to counteract its effect). Lastly, even if the DMIs are resolved in opposite directions, the different derived alleles will be at low frequency when their interacting partner is lost, and therefore more likely to be lost by drift subsequently. Overall, this implies that hybrid speciation via reciprocal sorting is on average less likely than illustrated here. Furthermore, one-time secondary contact between two diverged populations (or species) is usually geographically restricted, and therefore tends to happen for small populations. However, this apparent rarity of hybrid speciation can be counteracted by the frequent formation of hybrid populations; this could suggest that

the reported cases of homoploid speciation may simply reflect a geographical distribution conducive to the formation and isolation of hybrid population. The Italian Sparrow seems to fit this scenario remarkably well [48].

The search for signs of hybrid speciation in very large populations, for example yeast, will be an exciting avenue for hybrid speciation research in the future (see e.g., [53]). That is because, firstly, multiple DMIs, both inter- and intraspecific, have been identified in yeast [30, 54, 55]. In addition, yeast can be easily maintained in the lab and lends itself to powerful experimental approaches. The large population size and the possibility to create and maintain multiple replicates are two strong advantages of experimental-evolution approaches to potentially quantify the reciprocal sorting of DMIs. Note that a case of natural homoploid hybrid speciation in yeast was recently reported, in which reproductive isolation was indeed attributed to postzygotic mechanisms [56].

One remarkable genetic incompatibility in yeast that could serve as a contributor to hybrid speciation was identified by Bui *et al.* [57]. Intriguingly, the phenotype of the incompatibility is an increased mutation rate, which could provide the fuel for adaptation to a new environment upon hybridization [58]. Intrinsic incompatibilities are usually deleterious, however in this specific case they could give a temporary advantage to the hybrid population. Following the formation of a hybrid population, the incompatible haplotype could spread in the population while the hybrid population is adapting to new (harsh) conditions. Such a mechanism could therefore be powerful in a colonization scenario. In the initial stage, the mutator phenotype might increase the possibility of recruiting adaptive mutations. Sorting of the incompatibility happens only at a later stage, which results in a synergy of forming genetic isolating barriers from the parental source populations while delaying the incompatibility-sorting phase until the population has adapted to its environment.

Conclusion

The probability of hybrid speciation is subject to continuing debate [9, 10, 24]. The reciprocal sorting of parental incompatibilities has been proposed as one credible mechanism to achieve hybrid speciation. Our work legitimates the existing disagreement by demonstrating that the hybrid speciation probability via reciprocal parental incompatibility sorting is highly variable and dependent on the linkage architecture and the dominance type of the involved incompatibilities. Specifically, the linkage architecture determines not only whether hybrid speciation is achievable or not, but also whether equal or unequal initial proportions of the parental populations are favorable for hybrid speciation. In addition, we show that across all studied scenarios, intermediate recombination rates maximize the likelihood of reciprocal sorting; i.e., interactions on the same chromosome are favorable for hybrid speciation. Altogether, we arrive at the prediction that in nature, hybrid speciation via reciprocal sorting of incompatibilities should indeed be rare; at the same time however, it can become almost deterministic (and, thus, repeatable) under specific genetic and demographic circumstances. Such circumstances could be met in microorganisms such as yeast, which have large population sizes and potential for repeated hybridization.

Supporting information

S1 Fig. Fitness landscapes for the codominant and recessive models. Here, we assume the “Adjacent ABAB” architecture and drop the indices to improve readability of the figure. The white dots indicate the position of the parental genotypes and the black dots the position of the hybrid speciation haplotypes. We use the “default” set of parameters: $\alpha_i = \beta_i = 0.001$, $\epsilon = -0.2$. The fitness advantage of the parental genotypes as compared to the ancestral genotype, chosen

here as $0.001^4 \approx 0.004$, is too small to be visible in the illustration. The arrangement of the haplotypes is arbitrary.

(TIF)

S2 Fig. Correlation between fixation of allele A_1 and allele A_2 for codominant (left) and recessive (right) DMIs. Each incompatibility pair is located on a different chromosome.

Colors indicate recombination rates between the A_k and B_k loci involved in the DMIs: $r = 0.5$ in black, $r = 0.1$ in purple, $r = 0.05$ in blue and $r = 0.005$ in cyan. Other parameters used are: $\alpha_i = \beta_j = 0.001$, $\epsilon = -0.2$, $N = 5000$.

(TIF)

S3 Fig. Time to resolution of both DMIs conditioning on hybrid speciation. Parameters used here are: $\alpha_i = \beta_j = 0.001$, $\epsilon = -0.2$, $N = 5000$. For any initial frequency, we performed 1000 simulations and then extracted those simulations that resulted in hybrid speciation.

(TIF)

S4 Fig. Marginal fitness of the four derived alleles in the deterministic model (no drift) for the first 100 generations. For the “Adjacent ABAB” and “Crossed AABB” architectures, alleles A_2 and B_1 have a marginal fitness advantage over A_1 and B_2 . For the “Adjacent ABBA” and “Crossed ABBA” architectures, alleles A_1 and A_2 have a marginal fitness advantage over B_1 and B_2 . Lastly, for the “Nested ABAB” and “Nested AABB” architectures, alleles A_2 and B_2 have a marginal fitness advantage over A_1 and B_1 . All DMIs represented here are codominant. Both parental populations contributed equally to the hybrid population, $i_p = 0.5$. All loci are equidistant with a recombination rate between adjacent loci of $r = 0.2$. Others parameters used are:

$\alpha_i = \beta_j = 0.001$, $\epsilon = -0.2$.

(TIF)

S5 Fig. Relative fixation probability of an unlinked neutral marker introduced by a parental individual migrating into the monomorphic hybrid population. This probability is displayed relative to the fixation probability of a similar neutral marker that appears in a hybrid individual, $p_n = 1/5000$. The x-axis corresponds to the recombination rate between the different incompatible loci; the neutral marker is always located on a different chromosome. We estimated the fixation probabilities over 10^6 replicates. Black lines indicate when the fixation probability is significantly distinct from $p_n = 1/5000$: any data point below the solid lines is significantly different under a Bonferroni correction (i.e. $\frac{p}{p_n} < 0.725$ or $\frac{p}{p_n} > 1.305$). Dashed lines correspond to 95% confidence intervals without Bonferroni correction and provide a visual guide.

(TIF)

S6 Fig. Illustration of the fitness landscapes for a single DMI in our model and an “adaptive” DMI from [24], S1 Fig. To facilitate the comparison, we use the notation of [24], as presented in Table S1 Table. Parametrization was chosen such that the fitness differences between parental and ancestral genotypes and the strength of the incompatibility for the homozygote incompatible ABAB are the same.

(TIF)

S7 Fig. Probability of fixation of allele A following formation of the hybrid population. In one case (red dots), haplotype Ab is introduced at proportion i_p and aB at proportion $1 - i_p$. To demonstrate that the frequency-dependent effect is dependent on the relative proportion of the A and B allele, in the other case (black dots), both haplotypes Ab and aB are introduced at proportion i_q and the ancestral haplotype ab is introduced at proportion $1 - 2i_q$.

(TIF)

S8 Fig. Dominance affects the probability of recovering the ancestral haplotype in a hybrid population. Panel (a) shows the probability of recovering the ancestral haplotype ab in an isolated hybrid population for the two dominance schemes both for neutral ($\alpha = \beta = 0$) and slightly advantageous mutations ($\alpha = \beta = 0.001$). Panel (b) tracks the marginal fitness of the ab haplotype in the deterministic model, for an (almost) symmetric contact (solid lines) and an asymmetric proportion of the parental genomes (dashed lines). Blue and red dots correspond to the data from panel (a). Due to the masking effect of recessivity, the marginal fitness of ab is always lower in the recessive case than in the codominant case. Panels (c) and (d) correspond to subsets of panel (b) and illustrate at which point the order of probabilities is reversed for symmetric (panel (d)) and asymmetric contact (panel (c)). Order inversion means that from this time point onwards, masking no longer provides an advantage to the derived alleles. The longer the derived alleles are masked by recessivity, the more likely the ancestral haplotype will fix while both derived alleles are present at an equally low frequency, and therefore susceptible to be lost through drift. Other parameters used are: $N = 5000$, $\epsilon = -0.2$ and $\alpha = \beta = 0.001$ for panels (b-d). In panel (a), each data point is obtained from 2000 simulations.

(TIF)

S9 Fig. Probability of recovering the ancestral haplotype increases with the time it takes to resolve the DMIs. Colors represent different recombination rates between the A and B loci. Each dot is the result of 1000 simulations. Parameters used are: $\alpha = \beta = 0.001$ and $N = 5000$.

(TIF)

S10 Fig. Hybrid speciation probability for recessive DMIs as a function of recombination for the six possible linkage architectures. We represent this probability for four different population sizes: $N = 50$ in yellow, $N = 500$ in orange, $N = 5000$ in red and $N = 50000$ in purple. Other parameters are $\alpha_k = \beta_j = 0.001$, $\epsilon = -0.2$ and $i_p = 0.5$ (i.e. the contribution of both parental populations is symmetric). This figure corresponds to Fig 5 for codominant DMIs.

(TIF)

S11 Fig. Proportions of possible evolutionary outcomes for the “Crossed ABBA” architecture. The x-axis corresponds to the recombination rate and each panel shows a different population size. To better illustrate the underlying mechanisms, we represent both hybrid haplotypes separately with $A_1B_2b_1a_2$ in green and $a_1b_2B_1A_2$ in yellow.

(TIF)

S12 Fig. Proportions of possible evolutionary outcomes for the “Adjacent ABBA” architecture. The x-axis corresponds to the recombination rate and each panel shows a different population size. To better illustrate the underlying mechanisms, we represent both hybrid haplotypes separately with $A_1B_2b_1a_2$ in green and $a_1b_2B_1A_2$ in yellow.

(TIF)

S13 Fig. Hybrid speciation probability as a function of the population size for recessive (top) and codominant (bottom) DMIs. We focus on the two architectures that displayed a second local maximum of hybrid speciation probability for low recombination rate, here $r = 0.005$. Each color corresponds to a different evolutionary outcome.

(TIF)

S14 Fig. Hybrid speciation probability for (quasi-)lethal DMIs as a function of recombination. We consider both codominant and recessive DMIs and two population sizes. Hybrid speciation does not occur for codominant DMIs regardless of population size. For recessive

DMIs, the hybrid speciation probability is qualitatively identical to the less deleterious case, see [S10 Fig](#) with the blue dots here matching the red dots in [S10 Fig](#) and the red diamonds the purple dots in [S10 Fig](#). Parameters used are: $\alpha = \beta = 0.001$, $\epsilon = -0.99$.

(TIF)

S15 Fig. Time to hybrid speciation is relatively unaffected by the dominance of the epistatic interaction, the strength of the epistasis or the linkage architecture (here: Adjacent ABAB, Adjacent ABBA, Crossed ABAB). Each row of panels represents a different linkage architecture, for which the time to hybrid speciation is shown for codominant (left), recessive (middle), and recessive lethal DMIs (right). We show the average time to hybrid speciation, i.e. to fixation of one of the two hybrid haplotypes, scaled by the size of the population as a function of the recombination rate. Colors indicate different population sizes; purple: $N = 50000$, red: $N = 5000$, orange: $N = 500$ yellow: $N = 50$. Each set of simulations was obtained from 1000 simulations, of which those resulting in hybrid speciation were retained. We only show the time to hybridization if we observed at least 4 occurrences of hybrid speciation. Standard errors are represented by black bars.

(TIF)

S16 Fig. Time to hybrid speciation is relatively unaffected by the dominance of the epistatic interaction, the strength of the epistasis or the linkage architecture (here: Crossed ABBA, Nested ABAB, Nested AABB). Each row of panels represents a different linkage architecture, for which the time to hybrid speciation is shown for codominant (left), recessive (middle), and recessive lethal DMIs (right). We show the average time to hybrid speciation, i.e. to fixation of one of the two hybrid haplotypes, scaled by the size of the population as a function of the recombination rate. Colors indicate different population sizes; purple: $N = 50000$, red: $N = 5000$, orange: $N = 500$ yellow: $N = 50$. Each set of simulations was obtained from 1000 simulations, of which those resulting in hybrid speciation were retained. We only show the time to hybridization if we observed at least 4 occurrences of hybrid speciation. Standard errors are represented by black bars.

(TIF)

S17 Fig. The time to hybrid speciation is similar between codominant and recessive DMIs despite a much faster resolution of the two DMIs in the codominant case. We show the time to resolution of the two DMIs (blue and cyan) and of loss of all polymorphism (i.e., a haplotype has fixed; black and gray). We compare the time of fixation of a hybrid haplotype (black) to the average fixation time of a haplotype (gray). In addition, we compare the average time to resolution of both DMIs for all evolutionary outcomes (cyan) and conditioned on the occurrence of hybrid speciation (blue). The linkage architecture used here is “Adjacent ABAB”. Each parameter set was obtained from 1000 simulations. We only display the time to hybrid speciation if we observed at least 4 occurrences of hybrid speciation for the respective parameter combination.

(TIF)

S1 Table. Comparison of the parametrization of genetic incompatibilities between the model of [24] (top) and our model (bottom). The derived alleles follow the same nomenclature. However, the ancestral alleles, which we called “a” and “b” are represented by “x” in the Schumer model.

(PDF)

S1 File. Supplementary information.

(PDF)

Acknowledgments

We thank Roger Butlin, Florian Clemente, Aaron Comeault, Ilse Höllinger, Molly Schumer and the members of the Bank lab for helpful discussion and comments on the manuscript. We thank two anonymous reviewers for their careful reading and insightful comments on the manuscript. This research was supported by the Fundação Calouste Gulbenkian.

Author Contributions

Conceptualization: Alexandre Blanckaert, Claudia Bank.

Investigation: Alexandre Blanckaert.

Software: Alexandre Blanckaert.

Supervision: Claudia Bank.

Visualization: Alexandre Blanckaert.

Writing – original draft: Alexandre Blanckaert, Claudia Bank.

Writing – review & editing: Alexandre Blanckaert, Claudia Bank.

References

1. Barton N, Bengtsson B. The barrier to genetic exchange between hybridising populations. *Heredity*. 1986; 57(3):357–376. <https://doi.org/10.1038/hdy.1986.135> PMID: 3804765
2. Rieseberg L. Hybrid Origins of Plant Species. *Annual Review of Ecology and Systematics*. 1997; 28(2):359–389. <https://doi.org/10.1146/annurev.ecolsys.28.1.359>
3. Arnold M, Bulger M, Burke J, Hempel A, Williams J. Natural hybridization: how low can you go and still be important? *Ecology*. 1999; 80(2):371–381.
4. Buerkle C, Morris R, Asmussen M, Rieseberg L. The likelihood of homoploid hybrid speciation. *Heredity*. 2000; 84(4):441–451. <https://doi.org/10.1046/j.1365-2540.2000.00680.x> PMID: 10849068
5. Barton N. The role of hybridization in evolution. *Molecular Ecology*. 2001; 10(3):551–568. <https://doi.org/10.1046/j.1365-294x.2001.01216.x> PMID: 11298968
6. Mallet J. Hybrid speciation. *Nature*. 2007; 446(7133):279–283. <https://doi.org/10.1038/nature05706> PMID: 17361174
7. Abbott R, Albach D, Ansell S, Arntzen J, Baird S, Bierne N, et al. Hybridization and speciation. *Journal of Evolutionary Biology*. 2013; 26(2):229–246. <https://doi.org/10.1111/j.1420-9101.2012.02599.x> PMID: 23323997
8. Servedio M, Hermisson J, Doorn G. Hybridization may rarely promote speciation. *Journal of evolutionary biology*. 2013; 26(2):282–285. <https://doi.org/10.1111/jeb.12038> PMID: 23324008
9. Nieto Feliner G, Álvarez I, Fuertes-Aguilar J, Heuertz M, Marques I, Moharrekk F, et al. Is homoploid hybrid speciation that rare? An empiricist's view. *Heredity*. 2017; 118(6):513–516. <https://doi.org/10.1038/hdy.2017.7> PMID: 28295029
10. Schumer M, Rosenthal G, Andolfatto P. What do we mean when we talk about hybrid speciation? *Heredity*. 2018; p. 11–31.
11. Hedrick P. Adaptive introgression in animals: Examples and comparison to new mutation and standing variation as sources of adaptive variation. *Molecular Ecology*. 2013; 22(18):4606–4618. <https://doi.org/10.1111/mec.12415> PMID: 23906376
12. Seehausen O. Conditions when hybridization might predispose populations for adaptive radiation. *Journal of Evolutionary Biology*. 2013; 26(2):279–281. <https://doi.org/10.1111/jeb.12026> PMID: 23324007
13. Seehausen O, Takimoto G, Roy D, Jokela J. Speciation reversal and biodiversity dynamics with hybridization in changing environments. *Molecular Ecology*. 2008; 17(1):30–44. <https://doi.org/10.1111/j.1365-294X.2007.03529.x> PMID: 18034800
14. Turissini D, McGirr J, Patel S, David J, Matute D. The rate of evolution of postmating-prezygotic reproductive isolation in *Drosophila*. *Molecular Biology and Evolution*. 2017; 35(2):312–334. <https://doi.org/10.1093/molbev/msx271>

15. Todesco M, Pascual M, Owens G, Ostevik K, Moyers B, Hübner S, et al. Hybridization and extinction. *Evolutionary Applications*. 2016; 9(7):892–908. <https://doi.org/10.1111/eva.12367> PMID: 27468307
16. Schwarz D, Matta B, Shakir-Botteri N, McPherson B. Host shift to an invasive plant triggers rapid animal hybrid speciation. *Nature*. 2005; 436(7050):546–549. <https://doi.org/10.1038/nature03800> PMID: 16049486
17. Mavárez J, Salazar C, Bermingham E, Salcedo C, Jiggins C, Linares M. Speciation by hybridization in *Heliconius* butterflies. *Nature*. 2006; 441(7095):868–871. <https://doi.org/10.1038/nature04738> PMID: 16778888
18. Larsen P, Marchan-Rivadeneira M, Baker R. Natural hybridization generates mammalian lineage with species characteristics. *Proceedings of the National Academy of Sciences*. 2010; 107(25):11447–11452. <https://doi.org/10.1073/pnas.1000133107>
19. Hermansen J, Sæther S, Elgvin T, Borge T, Hjelle E, Sætre G. Hybrid speciation in sparrows I: Phenotypic intermediacy, genetic admixture and barriers to gene flow. *Molecular Ecology*. 2011; 20(18):3812–3822. <https://doi.org/10.1111/j.1365-294X.2011.05183.x> PMID: 21771138
20. Kang J, Scharl M, Walter R, Meyer A. Comprehensive phylogenetic analysis of all species of swordtails and platies (Pisces: *Genus Xiphophorus*) uncovers a hybrid origin of a swordtail fish, *Xiphophorus monticolus*, and demonstrates that the sexually selected sword originated in the ancestral lineage of the genus, but was lost again secondarily. *BMC Evolutionary Biology*. 2013; 13(1). <https://doi.org/10.1186/1471-2148-13-25> PMID: 23360326
21. Yakimowski S, Rieseberg L. The role of homoploid hybridization in evolution: A century of studies synthesizing genetics and ecology. *American Journal of Botany*. 2014; 101(8):1247–1258. <https://doi.org/10.3732/ajb.1400201> PMID: 25156978
22. Lamichhaney S, Han F, Webster M, Andersson L, Grant B, Grant P. Rapid hybrid speciation in Darwin's finches. *Science*. 2018; 359(6372):224–228. <https://doi.org/10.1126/science.aao4593> PMID: 29170277
23. Schumer M, Rosenthal G, Andolfatto P. HOW COMMON IS HOMOPLOID HYBRID SPECIATION? *Evolution*. 2014; 68(6):1553–1560. <https://doi.org/10.1111/evo.12399> PMID: 24620775
24. Schumer M, Cui R, Rosenthal G, Andolfatto P. Reproductive isolation of hybrid populations driven by genetic incompatibilities. *PLoS Genet*. 2015; 11(3):e1005041. <https://doi.org/10.1371/journal.pgen.1005041> PMID: 25768654
25. Comeault A. The genomic and ecological context of hybridization affects the probability that symmetrical incompatibilities drive hybrid speciation. *Ecology and evolution*. 2018; 8(5):2926–2937. <https://doi.org/10.1002/ece3.3872> PMID: 29531706
26. Bateson W. Heredity and variation in modern lights. *Darwin and modern science*. 1909; p. 85–101.
27. Dobzhansky T. Studies on hybrid sterility. II. Localization of sterility factors in *Drosophila pseudoobscura* hybrids. *Genetics*. 1936; 21(2):113. PMID: 17246786
28. Muller H. Isolating mechanisms, evolution and temperature. In: *Biol. Symp.* vol. 6; 1942. p. 71–125.
29. Presgraves D. A fine-scale genetic analysis of hybrid incompatibilities in *Drosophila*. *Genetics*. 2003; 163(3):955–972. PMID: 12663535
30. Kao KC, Schwartz K, Sherlock G. 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*. 2010; 6(7):e1001038. <https://doi.org/10.1371/journal.pgen.1001038> PMID: 20686707
31. Corbett-Detig R, Zhou J, Clark A, Hartl D, Ayroles J. Genetic incompatibilities are widespread within species. *Nature*. 2013; 504(7478):135–137. <https://doi.org/10.1038/nature12678> PMID: 24196712
32. Scarpino S, Hunt P, Garcia-De-leon F, Juenger T, Scharl M, Kirkpatrick M. Evolution of a genetic incompatibility in the genus *xiphophorus*. *Molecular Biology and Evolution*. 2013; 30(10):2302–2310. <https://doi.org/10.1093/molbev/mst127> PMID: 23894140
33. Colomé-Tatché M, Johannes F. Signatures of Dobzhansky–Muller Incompatibilities in the Genomes of Recombinant Inbred Lines. *Genetics*. 2016; 202(2):825–841. <https://doi.org/10.1534/genetics.115.179473> PMID: 26680662
34. Ono J, Gerstein A, Otto S. Widespread Genetic Incompatibilities between First-Step Mutations during Parallel Adaptation of *Saccharomyces cerevisiae* to a Common Environment. *PLoS biology*. 2017; 15(1):e1002591. <https://doi.org/10.1371/journal.pbio.1002591> PMID: 28114370
35. Marsit S, Leducq J, Durand E, Marchant A, Filteau M, et al. Evolutionary biology through the lens of budding yeast comparative genomics. *Nature Reviews Genetics*. 2017; 18(10):581–598. <https://doi.org/10.1038/nrg.2017.49> PMID: 28714481
36. Guerrero R, Muir C, Josway S, Moyle L. Pervasive antagonistic interactions among hybrid incompatibility loci. *PLoS Genetics*. 2017; 13(6):e1006817. <https://doi.org/10.1371/journal.pgen.1006817> PMID: 28604770

37. Orr H. The population genetics of speciation: the evolution of hybrid incompatibilities. *Genetics*. 1995; 139(4):1805–1813. PMID: [7789779](#)
38. Gross B, Rieseberg L. The ecological genetics of homoploid hybrid speciation. *Journal of heredity*. 2004; 96(3):241–252. <https://doi.org/10.1093/jhered/esi026> PMID: [15618301](#)
39. Turelli M, Orr H. Dominance, epistasis and the genetics of postzygotic isolation. *Genetics*. 2000; 154(4):1663–79. PMID: [10747061](#)
40. Bank C, Bürger R, Hermisson J. The Limits to Parapatric Speciation: Dobzhansky–Muller Incompatibilities in a Continent–Island Model. *Genetics*. 2012; 191(3):845–863. <https://doi.org/10.1534/genetics.111.137513> PMID: [22542972](#)
41. Haldane JBS. A mathematical theory of natural and artificial selection, part V: selection and mutation. In: *Mathematical Proceedings of the Cambridge Philosophical Society*. vol. 23. Cambridge Univ Press; 1927. p. 838–844.
42. Kimura M. On the probability of fixation of mutant genes in a population. *Genetics*. 1962; 47(6):713. PMID: [14456043](#)
43. Gross B. Genetic and phenotypic divergence of homoploid hybrid species from parental species. *Heredity*. 2012; 108(3):157. <https://doi.org/10.1038/hdy.2011.80> PMID: [21915149](#)
44. Via S. Divergence hitchhiking and the spread of genomic isolation during ecological speciation-with-gene-flow. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 2012; 367(1587):451–460. <https://doi.org/10.1098/rstb.2011.0260>
45. Feder J, Egan S, Nosil P. The genomics of speciation-with-gene-flow. *Trends in Genetics*. 2012; 28(7):342–350. <https://doi.org/10.1016/j.tig.2012.03.009> PMID: [22520730](#)
46. Wu CI. The genic view of the process of speciation. *Journal of Evolutionary Biology*. 2001; 14(6):851–865. <https://doi.org/10.1046/j.1420-9101.2001.00335.x>
47. Yeaman S. Genomic rearrangements and the evolution of clusters of locally adaptive loci. *Proceedings of the National Academy of Sciences*. 2013; 110(19):E1743–E1751. <https://doi.org/10.1073/pnas.1219381110>
48. Runemark A, Trier C, Eroukhmanoff F, Hermansen J, Matschiner M, Ravinet M, et al. Variation and constraints in hybrid genome formation. *Nature Ecology & Evolution*. 2018;. <https://doi.org/10.1038/s41559-017-0437-7>
49. Orr H. A mathematical model of Haldane's rule. *Evolution*. 1993; 47(5):1606–1611. <https://doi.org/10.1111/j.1558-5646.1993.tb02179.x> PMID: [28564905](#)
50. Coyne J, Orr H. *Speciation*. Sinauer Associates Sunderland, MA; 2004.
51. Höllinger I, Hermisson J. Bounds to parapatric speciation: A Dobzhansky-Muller incompatibility model involving autosomes, X chromosomes, and mitochondria. *Evolution*. 2017; 71(5):1366–1380. <https://doi.org/10.1111/evo.13223> PMID: [28272742](#)
52. Blanckaert A, Hermisson J. The limits to parapatric speciation II: Strengthening a preexisting genetic barrier to gene flow in parapatry. *Genetics*. 2018; 209(1):241–254. <https://doi.org/10.1534/genetics.117.300652> PMID: [29496748](#)
53. Greig D, Louis E, Borts R, Travisano M. Hybrid speciation in experimental populations of yeast. *Science*. 2002; 298(5599):1773–1775. <https://doi.org/10.1126/science.1076374> PMID: [12459586](#)
54. Matute D, Butler I, Turissini D, Coyne J. A Test of the Snowball Theory for the Rate of Evolution of Hybrid Incompatibilities. *Source: Science, New Series*. 2010; 329(5998):1518–1521.
55. Hou J, Schacherer J. Negative epistasis: a route to intraspecific reproductive isolation in yeast? *Current genetics*. 2016; 62(1):25–29. <https://doi.org/10.1007/s00294-015-0505-y> PMID: [26164016](#)
56. Leducq J, Nielly-Thibault L, Charron G, Eberlein C, Verta J, Samani P, et al. Speciation driven by hybridization and chromosomal plasticity in a wild yeast. *Nature Microbiology*. 2016; 1(1):15003. <https://doi.org/10.1038/nmicrobiol.2015.3> PMID: [27571751](#)
57. Bui D, Dine E, Anderson J, Aquadro C, Alani E. A genetic incompatibility accelerates adaptation in yeast. *PLoS genetics*. 2015; 11(7):e1005407. <https://doi.org/10.1371/journal.pgen.1005407> PMID: [26230253](#)
58. Taddei F, Radman M, Maynard-Smith J, Toupance B, Gouyon P, Godelle B. Role of mutator alleles in adaptive evolution. *Nature*. 1997; 387(6634):700. <https://doi.org/10.1038/42696> PMID: [9192893](#)