

Title: Reproductive isolation of hybrid populations driven by genetic incompatibilities

Molly Schumer^a, Rongfeng Cui^{c,d}, Gil G. Rosenthal^{c,d} and Peter Andolfatto^{a,b}

^aDepartment of Ecology and Evolutionary Biology and the ^bLewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ 08544, USA.

^cDepartment of Biology, Texas A&M University, TAMU, College Station, TX, USA

^d Centro de Investigaciones Científicas de las Huastecas “Aguazarca”, Calnali, Hidalgo, Mexico

Author contact information:

Corresponding author: Molly Schumer: schumer@princeton.edu, 106A Guyot Hall, Princeton University, Princeton NJ, 08544. Tel: 609-258-3160.

Rongfeng Cui: rcui@mail.bio.tamu.edu, 3258 Texas A&M University, College Station, TX 77843. Tel: 979-845-3614.

Gil Rosenthal: groenthal@bio.tamu.edu, 3258 Texas A&M University, College Station, TX 77843. Tel: 979-845-3614.

Peter Andolfatto: pandolfa@princeton.edu, 106A Guyot Hall, Princeton University, Princeton NJ, 08544. Tel. 609-258-3109.

Key terms: Hybrid speciation, hybrid incompatibilities, Bateson Dobzhansky Muller incompatibilities, reproductive isolation

Running title: A simple route to hybrid speciation?

1 **Abstract**

2 Despite its role in homogenizing populations, hybridization has also been
3 proposed as a means to generate new species. The conceptual basis for this idea is that
4 hybridization can result in novel phenotypes through recombination between the parental
5 genomes, allowing a hybrid population to occupy ecological niches unavailable to
6 parental species. A key feature of these models is that these novel phenotypes
7 ecologically isolate hybrid populations from parental populations, precipitating
8 speciation. Here we present an alternative model of the evolution of reproductive
9 isolation in hybrid populations that occurs as a simple consequence of selection against
10 incompatibilities. Unlike previous models, our model does not require small population
11 sizes, the availability of new niches for hybrids or ecological or sexual selection on
12 hybrid traits. We show that reproductive isolation between hybrids and parents evolves
13 frequently and rapidly under this model, even in the presence of substantial ongoing
14 migration with parental species and strong selection against hybrids. Our model predicts
15 that multiple distinct hybrid species can emerge from replicate hybrid populations formed
16 from the same parental species, potentially generating patterns of species diversity and
17 relatedness that resemble an adaptive radiation.

18

19 **Introduction**

20 The evolutionary significance of hybridization has been a hotly debated topic for
21 decades (1). Homoploid hybrid speciation, speciation that occurs as a result of
22 hybridization without a ploidy change (2), is generally thought to be an exceptionally rare
23 outcome of hybridization and there are only a handful of well-supported cases of this

1 phenomenon (3). Though it is not uncommon for species' genomes to exhibit evidence of
 2 past hybridization, hybrids are often thought to be weakly isolated from parental species,
 3 though few studies have explicitly investigated this. The most convincing cases of
 4 homoploid hybrid speciation to date have identified hybrid populations that have
 5 transgressive traits that differentiate their ecologically or sexually selected phenotypes
 6 from parental species, establishing reproductive isolation between hybrids and parents (4-
 7 7).

8 Partly as an outgrowth of this empirical work, recent theoretical work has focused
 9 on transgressive traits generated via hybridization and their role in facilitating
 10 reproductive isolation between parental and hybrid populations. These models have
 11 shown that hybrid speciation is a likely outcome if a new niche is available and
 12 transgressive traits allow hybrids to colonize that niche (8, 9; discussed in 10, 11).
 13 However, it is generally difficult to evaluate the likelihood of generating novel traits via
 14 hybridization (but see 12) or the likelihood that the trait will allow hybrids to colonize a
 15 new environment, making it challenging to predict the importance of hybridization in
 16 generating new species. Other models of homoploid hybrid speciation have considered
 17 the potential for combined effects of inbreeding and chromosomal rearrangements to
 18 generate reproductively isolated hybrid species (13, 14), and empirical results from
 19 sunflowers provide support for this mechanism (15).

20 We present a different but simple model in which reproductive isolation between
 21 hybrid and parental populations emerges as a consequence of selection on genetic
 22 incompatibilities in a hybrid swarm. The first genetic model of speciation, described by
 23 Bateson, Dobzhansky and Muller (the BDM model, Fig. S1, 16-18), predicts that

1 mutations at two genetic loci differentially accumulating along two lineages can
2 negatively interact in their hybrids. The genetic incompatibility of hybrids constitutes a
3 key component of reproductive isolation between many species, and is the basis for the
4 biological species concept (18). Empirical research has suggested that these types of
5 interactions are remarkably common (19-23; reviewed in 18, 24, 25). Though the theory
6 of BDM incompatibilities was originally formulated in the context of allopatrically
7 diverging species, our model investigates the fate of these incompatibilities in hybrid
8 populations.

9 Under the simplest BDM scenario, derived genotypes are neutral, meaning that
10 they have the same fitness as ancestral genotypes. Following hybridization, selection on
11 these incompatibilities in a hybrid swarm is expected to result in the fixation of genotype
12 combinations that are not incompatible with either parental species. However,
13 incompatibilities may also frequently arise due to adaptive evolution within lineages or as
14 coevolving pairs of loci (Fig. S1, S2, 18, 26-28). Here we show that selection on these
15 types of incompatibility pairs can result in the fixation of genotype combinations that
16 contribute to isolation between the hybrid population and one or the other parental
17 species. In the presence of multiple pairs of such incompatibilities, this process can result
18 in the rapid evolution of reproductive isolation of hybrid populations from both parental
19 species.

20 Two features of this model make it particularly plausible biologically. First,
21 empirical studies have revealed that hybrid incompatibilities rapidly accumulate between
22 species (24), providing ample raw material for our model to work. Second, hybrid
23 populations in which hybrids outnumber parental species occur frequently in nature,

1 either in stable hybrid zones (29, 30) or as a result of environmental disturbance (31, 32).
 2 These characteristics suggest that the evolution of reproductive isolation of hybrid
 3 populations by this mechanism, a first step towards homoploid hybrid speciation, could
 4 be a common evolutionary outcome of hybridization.

6 **Results**

8 *Modeling selection against hybrid incompatibilities*

9 In the simplest model of a hybrid population, a mixture of individuals from both
 10 parental species form a new isolated population and mate randomly with respect to
 11 genotype (Fig. 1A). Using theory developed by Karlin (33), one can model the effect of
 12 selection at two polymorphic loci on gamete frequencies of a diploid sexual population
 13 (see Methods and Supplement 1). Using this two-locus selection model, one can see that
 14 the direction of fixation depends on the initial frequency of the parental alleles (f , see Fig.
 15 1A and S3) and dominance at each locus (h , see Fig. S3).

16 This purely deterministic model of selection on hybrid incompatibilities is
 17 unrealistic because even large populations experience some degree of genetic drift. We
 18 thus extended the model to include genetic drift, which can affect the speed and direction
 19 of fixation of incompatibility pairs (Fig. S4). For coevolving or adaptive BDM
 20 incompatibilities (Supporting Information 1, Fig. S1, S2), this model predicts that at
 21 equal admixture proportions ($f = 0.5$), a single codominant incompatibility pair has a 50%
 22 chance of fixing for either parental allele combination (Fig. 2, Supporting Information 2).
 23 Interestingly, while genetic drift in very small populations could accomplish the same

1 thing (8), the process described here occurs rapidly in large populations and is driven by
 2 deterministic selection (Fig. 2). Given this result, it is clear that large hybrid populations
 3 with two or more hybrid incompatibilities could, in principal, fix for one parental
 4 genotype at one incompatibility pair and the other parental genotype at the other
 5 incompatibility pair (Fig. 1B). This outcome would result in reproductive isolation of the
 6 hybrid population from both parental species.

7

8 *Simulations of an isolated hybrid population*

9 To investigate the dynamics of multiple incompatibility pairs, we simulated a
 10 large, randomly-mating and spatially isolated hybrid population with two pairs of
 11 unlinked hybrid incompatibility loci (see Methods; Fig. 1A, setting $m_1=m_2=0$). The
 12 fitness scheme used is that of a coevolutionary incompatibility model (Fig. S2), assuming
 13 that incompatibilities are codominant (i.e. $h=0.5$), that fitness is symmetric with respect
 14 to the parental source of alleles (i.e. $w_{ij} = w_{ji}$) and that the cumulative fitness effects of
 15 multiple incompatibility pairs is multiplicative. If hybrid populations fixed for the parent
 16 1 genotypes at one incompatibility pair and the parent 2 genotypes at the other, we
 17 considered the hybrid population as having evolved reproductive isolation from both
 18 parents.

19 In simulations of this simple scenario, reproductive isolation between hybrid and
 20 parental populations evolved frequently and rapidly. For two incompatibility pairs with
 21 selection coefficients (s) of 0.1, $47\pm 2\%$ of simulated hybrid populations became isolated
 22 from both parental species within an average of ~ 150 generations. Exploring a range of s

1 (0.1-0.5, Fig. S5), initial admixture proportions ($f=0.3-0.7$, Fig. S6), and population sizes
 2 (100-10,000 diploids), we conclude that, unless fitness of hybrids is close to 0 or ancestry
 3 is highly skewed, reproductive isolation evolves rapidly and with surprisingly high
 4 probability ($27\pm 2\%$ to $44\pm 2\%$ of the time; on average within 55 ± 14 to 190 ± 43
 5 generations, see Supporting Information 3).

6

7 *The effect of dominance and asymmetry in selection intensity*

8 In the above simulations, we assume that selection on different hybrid
 9 incompatibility interactions is symmetrical ($s_1=s_2$, Fig. S2), but it is unlikely that selection
 10 is truly equal on different hybrid genotype combinations. When fitness is completely
 11 asymmetrical (i.e. s_1 or $s_2=0$, as for neutral BDM incompatibilities), only strong genetic
 12 drift can cause the fixation of genotype pairs that are incompatible with either parent
 13 (Fig. S8, Supporting Information 4) as selection cannot do so (see Fig. S1, S3, S7,
 14 Supporting Information 4). This reliance on genetic drift implies that this process will be
 15 slow unless an extreme bottleneck is invoked.

16 In contrast, the dynamics of BDM incompatibilities resulting from adaptation
 17 within parental lineages can be quite different (Fig. S1). Notably, while selection may
 18 also be highly asymmetric in such cases (34), derived alleles have higher fitness than
 19 ancestral alleles resulting in the fixation of genotype combinations that are incompatible
 20 with both parents. We find that isolation evolves with similar frequency under
 21 asymmetric selection as long as selection is strong relative to drift (Supporting

1 Information 3D), because even weak selection will prevent the fixation of the ancestral
2 genotype.

3 We arbitrarily simulate codominant hybrid incompatibilities, but Karlin's model
4 (Fig. S3) shows that patterns of fixation are different depending on the value of h . In
5 particular, when h is zero or unity, fixation is not strongly dependent on admixture
6 proportions (Fig. S3). When we simulate variation in dominance among incompatibility
7 interactions (see Supporting Information 3E), we find that reproductive isolation between
8 hybrid populations and parental species evolves with comparable frequency ($44-48 \pm 2\%$
9 vs $47 \pm 2\%$ under the codominant scenario).

10

11 *Increasing the number of hybrid incompatibilities*

12 Recent empirical studies have suggested that most species are distinguished by
13 multiple hybrid incompatibilities (24, 35-40). Theoretically, barring extinction of the
14 hybrid population, increasing the number of incompatibilities should increase the
15 probability that a hybrid population will evolve isolation from both parental species. In
16 order to illustrate this, we simulated 3-6 unlinked hybrid incompatibility pairs
17 (Supporting Information 5). As expected, increasing the number of hybrid
18 incompatibilities increases the probability that the hybrid population will be isolated from
19 each parent by at least one incompatibility ($>96\%$ with 6 incompatibility pairs, Fig. 4,
20 Fig. S5, Supporting Information 5B).

21 We assume in most of our simulations that loci involved in hybrid
22 incompatibilities are completely unlinked. As the number of incompatibilities increases,

1 this becomes increasingly less likely. Physical linkage between loci involved in different
 2 epistatic interactions can reduce the frequency at which hybrid populations evolve
 3 isolation because sites are more likely to fix for the same parent (Fig. S9, Supporting
 4 Information 5C). Interestingly, when hybrid incompatibility loci are linked to a neutral
 5 BDM incompatibility, this does not significantly lower the frequency at which hybrid
 6 populations evolve reproductive isolation (Supporting Information 5B). Furthermore,
 7 linkage between coevolving incompatibilities and neutral BDM incompatibilities can
 8 more frequently result in fixation of neutral BDM incompatibilities for a parental
 9 genotype ($21 \pm 2\%$), resulting in stronger isolation between hybrid and parental
 10 populations (Supporting Information 5B).

11 The above simulations focus on simple models of hybrid incompatibility that
 12 show this process can occur in principle. To capture more biological realism in the
 13 number and types of incompatibilities, we simulated 20 incompatibility pairs with
 14 randomly determined genomic position and dominance, drawing selection coefficients
 15 from an exponential distribution with a mean $s=0.05$ and introducing asymmetry in
 16 selection (see Supporting Information 5D). In these simulations 95% of populations
 17 developed isolation from both parents within 500 generations. On average, hybrid
 18 population first developed isolation from both parents after 180 generations and were
 19 isolated by 7 incompatibility pairs by 500 generations. Interestingly, because
 20 incompatibilities with the largest fitness effects tend to fix first, hybrid populations
 21 developed considerable reproductive isolation from parental species even before all
 22 incompatibilities were fixed in the population (Fig. 3, Fig. S10). Overall, our simulations
 23 suggest that rapid evolution of reproductive isolation of hybrid populations is likely when

1 parental species are separated by several hybrid incompatibilities with varying levels of
2 asymmetry, linkage and dominance.

3

4 *Simulations of hybrid populations with migration*

5 We model hybrid population formation as a discrete event, but many hybrid
6 populations experience continued gene flow with parental species. This process may
7 impede the evolution of reproductive isolation by preventing the fixation of genetic
8 incompatibilities. To evaluate this, we simulated hybridization scenarios with ongoing
9 migration (Fig. 1A, $4Nm=8-40$). Even with substantial continuing gene flow from
10 parental populations, hybrid populations evolved reproductive isolation from them at
11 high frequencies (i.e. $37\pm2\%$ of simulations with two incompatibilities, $s=0.1$ and
12 $4Nm=8$; Supporting Information 6).

13 Intuitively, as chance plays an important role in which incompatibility pairs fix,
14 independently formed hybrid populations from the same parental species can evolve
15 isolation from each other. To demonstrate this effect, we simulated two hybrid
16 populations formed from the same parental species (Fig. S12). In the absence of
17 migration, the two hybrid populations evolved isolation from each other frequently
18 ($52\pm5\%$, as expected given two hybrid incompatibilities, see Supporting Information 6C).
19 Remarkably, this outcome was still observed frequently with relatively high gene flow
20 between the two hybrid populations ($18\pm4\%$ with $4Nm=8$ and two hybrid
21 incompatibilities, Supporting Information 6C).

1 Discussion

2

3 We describe a new model of the evolution of reproductive isolation of hybrid
 4 populations, a first step towards hybrid speciation. Unlike previous models of hybrid
 5 speciation, our model does not rely on transgressive hybrid traits, the availability of new
 6 ecological niches, or small population sizes, but rather deterministic selection in large
 7 populations against hybrid incompatibilities. Reproductive isolation can evolve rapidly
 8 under our proposed model; with moderate selection (i.e. $s=0.1$) on a small number
 9 incompatibility pairs in an allopatric hybrid population, reproductive barriers from both
 10 parents emerge with close to 50% probability and on average within ~150 generations.
 11 Hybrid reproductive isolation also evolves frequently with substantial levels of ongoing
 12 migration between hybrids and parental species. The widespread existence of hybrid
 13 incompatibilities and the rapid emergence of reproductive isolation predicted by this
 14 model imply that hybrid speciation by this mechanism could be a common natural
 15 consequence of hybridization.

16 It is important to note several factors that may influence how common this mode
 17 of hybrid reproductive isolation is in natural populations. In our model, isolation between
 18 hybrids and parents is inherently weaker than between the two parents. Nonetheless, we
 19 propose that this mechanism may be a crucial first step in limiting gene flow between
 20 hybrids and parents, allowing for the development of other isolating mechanisms. For
 21 example, theoretical work predicts that reinforcement can develop even when selection is
 22 moderate on hybrids (e.g. $s \sim 0.3$, 41-44). We also note that our model only represents
 23 fitness in terms of genetic incompatibilities and that hybrid populations can have lower

1 fitness as a result of ecological or sexual selection. For example, when parental species
2 exert negative sexual selection against hybrids, hybrid populations are significantly more
3 likely to be outcompeted by parents (Supporting Information 6C). This highlights the fact
4 that the plausibility of this model depends on the biology of the species in question;
5 parental species of swordtail fish and pupfish do not discriminate strongly against hybrids
6 (45-47), while this is not the case in mice (48).

7 An additional consideration is that this model will only be relevant during a
8 particular window of divergence between species. When the fitness of hybrid populations
9 is low, they are more prone to extinction or displacement by parental populations
10 (Supporting Information 3). This suggests that hybrid speciation is most likely to occur in
11 a period of evolutionary divergence during which species have accumulated some hybrid
12 incompatibilities but have not diverged to the point to which hybrids are largely inviable.
13 The most detailed work characterizing genetic incompatibilities has been in *Drosophila*
14 species, where hybrids generally have substantially reduced fitness compared to parents
15 (36, 37, 49). Hybrids between several other species studied to date, however, are affected
16 by fewer incompatibilities or incompatibilities of weaker effects (20, 35, 40, 50-53). Such
17 species could be more likely to form hybrid populations, and should be the focus of
18 future empirical research.

19 It is also important to note that reduced frequency of reproductive isolation with
20 increasing selection on hybrids can be mediated by an increase in the total number of
21 hybrid incompatibilities. In our simulations, we see a positive relationship between the
22 number of interactions and the probability of developing reproductive isolation (Fig. 4)
23 and a negative relationship between the total strength of selection on hybrids and the

1 probability of developing reproductive isolation (Fig. S5). This tradeoff suggests that
 2 reproductive isolation can evolve between hybrid and parental populations even when the
 3 fitness of hybrids is low (as in Fig. 3, Fig. S5).

4 Our model requires that coevolving incompatibilities or BDM incompatibilities
 5 arising from adaptive evolution frequently occur between species. Accumulating
 6 evidence suggests that incompatibilities arising from coevolution may be common (24,
 7 27, 54-57). For example, in marine copepods, coevolution between cytochrome *c* and
 8 cytochrome *c* oxidase results in a reciprocal breakdown of protein function in hybrids
 9 (57). In addition, the fact that many known incompatibility genes involve sexual conflict,
 10 selfish genetic elements, or pathogen defense suggests an important role for coevolution
 11 in the origin of incompatibilities (27, 54, 58, 59). Our model also applies to BDM
 12 incompatibilities that arise due to within-lineage adaptation. It is currently unknown
 13 whether effectively neutral BDIs are more common than adaptive BDIs. Though
 14 many incompatibilities have asymmetric selection on different hybrid genotypes (e.g. 22,
 15 23, 60), neutrality has not been established in these cases. Anecdotal evidence supports
 16 the idea that adaptive incompatibilities are common since many hybrid incompatibilities
 17 show evidence of positive selection within lineages (61), but the relative frequency of
 18 adaptive and neutral BDIs awaits answers from further empirical research.

19 The patterns predicted by this model are testable with empirical approaches. A
 20 large number of studies have successfully mapped genetic incompatibilities
 21 distinguishing species (19, 20, 36, 37, 39, 53, 62). Ancestry at these sites can be
 22 determined in putative hybrid species, and the relative contribution of parental-derived
 23 incompatibilities to reproductive isolation can be determined experimentally. For some

1 species, it may be possible to evaluate the dynamics of incompatibilities relative to the
 2 genetic background in hybrid swarms (63). We predict that many hybrid populations
 3 exhibiting postzygotic isolation from parental species will have fixed incompatibility
 4 pairs for each parental species. Some proposed cases of hybrid speciation report reduced
 5 fitness of offspring between parental and hybrid species consistent with the mechanism
 6 described here (5, 15, 64, 65) and are promising cases for further empirical research.

7 An intriguing implication of our model is that independently formed hybrid
 8 populations between the same parental species can develop reproductive isolation from
 9 each other. The likelihood of this outcome increases with the number of incompatibilities.
 10 In sunflowers, empirical studies of ecologically-mediated hybrid speciation have
 11 identified multiple hybrid species derived from the same parental species (66). It is
 12 interesting to note that selection against hybrid incompatibilities could generate the same
 13 pattern in replicate hybrid populations. In fact, this mechanism could generate a species
 14 diversity pattern similar to that expected from an adaptive radiation, with multiple closely
 15 related species arising in a short evolutionary window. This finding is striking because
 16 our model does not invoke adaptation and suggests that non-adaptive processes (i.e.
 17 selection against incompatibilities) could also explain clusters of rapidly arising, closely
 18 related species.

19

1 **Methods**

2 *Mathematical model of selection on hybrid incompatibilities*

3 To characterize evolution at hybrid incompatibility loci in hybrid populations
4 without drift, we used the equations described by Karlin (33) to calculate changes in
5 allele frequency as a result of two-locus selection. The frequency of gamete i at
6 generation t is given by

$$7 \quad x_i(t) = x_i(t-1) \frac{w_{i*}}{\bar{w}} + \epsilon_i r D \frac{w_{14}}{\bar{w}}, \quad (1)$$

8 where $\epsilon_1 = \epsilon_4 = -\epsilon_2 = -\epsilon_3 = -1$ and the marginal fitness of allele i ,

$$9 \quad w_i^* = \sum_{j=1}^4 x_j w_{ij}, \quad (2)$$

10 the mean fitness of the population,

$$11 \quad \bar{w} = \sum_{i=1}^4 \sum_{j=1}^4 x_i x_j w_{ij}, \quad (3)$$

12 w_{14} is the fitness of a double heterozygote, r is the recombination rate and D is linkage
13 disequilibrium between the two loci. These equations assume random mating, non-
14 overlapping generations and that fitness depends only on two-locus genotype and not on
15 whether the chromosome was maternally or paternally inherited ($w_{ij}=w_{ji}$). To model
16 changes in allele frequencies over time we developed a custom R script (available from
17 Dryad, doi:XXX). Iterating through the change in allele frequencies each generation as a
18 result of selection gives the expected patterns of fixation at incompatibility loci without
19 genetic drift (Fig. S3).

20 Karlin's model of fixation of hybrid incompatibilities does not realistically predict
21 expected patterns in natural populations because even large populations will have some
22 level of genetic drift. To model this we added multinomial sampling of N diploid

1 individuals and recalculated allele frequencies each generation (available from Dryad,
2 [doi:XXX](#)). Patterns of fixation incorporating genetic drift through multinomial sampling
3 show similar dynamics to the model lacking genetic drift, with the exception of several
4 equilibrium states specific to the latter (see Fig. S4, Supporting Information 2).

6 *Description of simulation program*

7 In order to simulate multiple pairs of hybrid incompatibilities in a large number of
8 individuals and demographic scenarios, we developed a custom c++ program, called
9 `admix'em` (github: <https://github.com/melop/admixem>). The code allows one to
10 specify the number and length of chromosomes and the genomic locations of hybrid
11 incompatibilities and neutral markers. The current implementation assumes non-
12 overlapping generations and diploid sexual individuals. When modeling linkage, we
13 assume a uniform recombination rate and one recombination event per chromosome per
14 meiosis. Unless otherwise specified, we model all pairs of hybrid incompatibility loci as
15 unlinked and on separate chromosomes. As we are interested in short-term dynamics, we
16 do not implement mutation.

17 Selection coefficients are assigned to particular allelic combinations according to
18 a hybrid fitness matrix (see Fig. S1,S2). Based on each individual's genotype at hybrid
19 incompatibility loci, we calculate total individual fitness w , defined as the probability of
20 survival of that individual. Total fitness across multiple incompatibility pairs is assumed
21 to be multiplicative. Each female mates with one randomly selected male (but incorporate
22 assortative mating in Supporting Information 6C), and produces a Poisson distributed
23 number of offspring with a mean of 2. After selection, if the carrying capacity (N) is not

1 reached, additional offspring from the same mating events will be drawn from a Poisson
2 distribution with a new mean of (carrying capacity – current population size)/number of
3 females. This process is repeated until carrying capacity is reached or females have no
4 available gametes (set to a maximum of 10).

5 All reported results are based on 500 replicate simulations, which were conducted
6 for 2000 generations. In the majority of simulations (except Supporting Information
7 5B,C) the hybrid population is initially colonized by 500 individuals of each parental
8 species. Hybrid and parental populations were modeled as spatially distinct with
9 migration parameters between them; most simulations specified one hybrid population
10 formed between two parental populations (Fig. 1A) but we also simulated a stepping-
11 stone model and a model with multiple independently formed hybrid populations
12 (Supporting Information 6, Fig. S11, Fig. S12). Details on individual simulations and
13 results can be found in the supporting text. Hybrid populations are considered to have
14 evolved reproductive barriers from both parents if they fix at least one incompatibility for
15 each parent; the actual strength of selection against offspring between hybrids and parents
16 will depend on the selection coefficient and number of incompatibilities.

17

18 **Acknowledgements**

19 We would like to thank members of the Rosenthal and Andolfatto labs for helpful
20 discussion and Clair Han, Ying Zhen, and Yaniv Brandvain for comments on earlier
21 versions of this manuscript. This work was supported by an NSF GRFP (DGE0646086)
22 and NSF DDIG (DEB-1405232) to M.S and an NSF IOS-0923825 to GGR.

23

1 **Author contributions**

2 M.S. conceived of the project. M.S., R.C., G.G.R, and P.A. designed the project.

3 M.S., R.C and P.A. wrote scripts and ran simulations. R.C. wrote admix'em program.

4 M.S., R.C., G.G.R and P.A. analyzed results and wrote the manuscript.

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

References

2

- 3 1. Arnold ML (2006) *Evolution through genetic exchange* (Oxford University Press,
4 Oxford, U.K.).
- 5 2. Mallet J (2007) Hybrid speciation. *Nature* 446(7133):279-283.
- 6 3. Schumer M, Rosenthal GG, & Andolfatto P (2014) How common is homoploid
7 hybrid speciation? *Evolution* 68:1553–1560.
- 8 4. Rieseberg LH, Vanfossen C, & Desrochers AM (1995) Hybrid speciation
9 accompanied by genomic reorganization in wild sunflowers. *Nature*
10 375(6529):313-316.
- 11 5. Rieseberg LH, *et al.* (2003) Major ecological transitions in wild sunflowers
12 facilitated by hybridization. *Science* 301(5637):1211-1216.
- 13 6. Lexer C, Welch ME, Raymond O, & Rieseberg LH (2003) The origin of
14 ecological divergence in *Helianthus paradoxus* (Asteraceae): Selection on
15 transgressive characters in a novel hybrid habitat. *Evolution* 57(9):1989-2000.
- 16 7. Mavarez J, *et al.* (2006) Speciation by hybridization in *Heliconius* butterflies.
17 *Nature* 441(7095):868-871.
- 18 8. Buerkle CA, Morris RJ, Asmussen MA, & Rieseberg LH (2000) The likelihood
19 of homoploid hybrid speciation. *Heredity* 84(4):441-451.
- 20 9. Duenez-Guzman EA, Mavarez J, Vose MD, & Gavrillets S (2009) Case studies
21 and mathematical models of ecological speciation. 4. Hybrid speciation in
22 butterflies in a jungle. *Evolution* 63(10):2611-2626.

- 1 10. Mavarez J & Linares M (2008) Homoploid hybrid speciation in animals.
2 *Molecular Ecology* 17(19):4181-4185.
- 3 11. Jiggins CD, Salazar C, Linares M, & Mavarez J (2008) Hybrid trait speciation and
4 *Heliconius* butterflies. *Philosophical Transactions of the Royal Society B-*
5 *Biological Sciences* 363(1506):3047-3054.
- 6 12. Stelkens R & Seehausen O (2009) Genetic distance between species predicts
7 novel trait expression in their hybrids. *Evolution* 63(4):884-897.
- 8 13. Grant V (1971) *Plant Speciation* (Columbia University Press, New York).
- 9 14. McCarthy EM, Asmussen MA, & Anderson WW (1995) A theoretical assessment
10 of recombinational speciation. *Heredity* 74:502-509.
- 11 15. Lai Z, *et al.* (2005) Extensive chromosomal repatterning and the evolution of
12 sterility barriers in hybrid sunflower species. *Genetics* 171(1):291-303.
- 13 16. Dobzhansky T (1937) *Genetics and the origin of species* (Columbia University
14 Press, New York).
- 15 17. Muller HJ (1942) Isolating mechanisms, evolution and temperature. *Biol Symp*
16 6:71-125.
- 17 18. Coyne JA & Orr HA (2004) *Speciation* (Sinauer Associates, Sunderland, MA).
- 18 19. Fishman L & Willis JH (2001) Evidence for Dobzhansky-Muller incompatibilities
19 contributing to the sterility of hybrids between *Mimulus guttatus* and *M-nasutus*.
20 *Evolution* 55(10):1932-1942.
- 21 20. Sweigart AL, Fishman L, & Willis JH (2006) A simple genetic incompatibility
22 causes hybrid male sterility in *mimulus*. *Genetics* 172(4):2465-2479.

- 1 21. Brideau NJ, *et al.* (2006) Two Dobzhansky-Muller genes interact to cause hybrid
2 lethality in *Drosophila*. *Science* 314(5803):1292-1295.
- 3 22. Meiklejohn CD, *et al.* (2013) An Incompatibility between a Mitochondrial tRNA
4 and Its Nuclear-Encoded tRNA Synthetase Compromises Development and
5 Fitness in *Drosophila*. *PLoS Genet* 9(1).
- 6 23. Sawamura K & Yamamoto MT (1997) Characterization of a reproductive
7 isolation gene, zygotic hybrid rescue, of *Drosophila melanogaster* by using
8 minichromosomes. *Heredity* 79:97-103.
- 9 24. Presgraves DC (2010) The molecular evolutionary basis of species formation. *Nat*
10 *Rev Genet* 11(3):175-180.
- 11 25. Orr HA & Presgraves DC (2000) Speciation by postzygotic isolation: forces,
12 genes and molecules. *Bioessays* 22(12):1085-1094.
- 13 26. Seehausen O, *et al.* (2014) Genomics and the origin of species. *Nat Rev Genet*
14 15(3):176-192.
- 15 27. Rieseberg LH & Blackman BK (2010) Speciation genes in plants. *Ann Bot*
16 106(3):439-455.
- 17 28. Johnson NA, Wolf JB, Brodie ED, III, & Wade MJ (2000) Gene interactions and
18 the origin of species. *Epistasis and the evolutionary process.*, eds Wolf J, Brodie
19 EI, & Wade M (Oxford University Press, New York), pp 197-212.
- 20 29. Sites JW (1994) Hybrid zones and the evolutionary process. *Science*
21 264(5159):727-728.

- 1 30. Ording GJ, Mercader RJ, Aardema ML, & Scriber JM (2010) Allochronic
2 isolation and incipient hybrid speciation in tiger swallowtail butterflies. *Oecologia*
3 162(2):523-531.
- 4 31. Fisher HS, Wong BBM, & Rosenthal GG (2006) Alteration of the chemical
5 environment disrupts communication in a freshwater fish. *Proc R Soc London Ser*
6 *B* 273(1591):1187-1193.
- 7 32. Hasselman DJ, *et al.* (2014) Human disturbance causes the formation of a hybrid
8 swarm between two naturally sympatric fish species. *Mol Ecol* 23:1137–1152.
- 9 33. Karlin S (1975) General 2-locus selection models - Some objectives, results and
10 interpretations. *Theor Popul Biol* 7(3):364-398.
- 11 34. Gavrillets S (1997) Hybrid zones with Dobzhansky-type epistatic selection.
12 *Evolution* 51(4):1027-1035.
- 13 35. Moyle LC & Graham EB (2006) Genome-wide associations between hybrid
14 sterility QTL and marker transmission ratio distortion. *Mol Biol Evol* 23(5):973-
15 980.
- 16 36. Masly JP & Presgraves DC (2007) High-resolution genome-wide dissection of the
17 two rules of speciation in *Drosophila*. *PLoS Biol* 5(9):1890-1898.
- 18 37. Presgraves DC (2003) A fine-scale genetic analysis of hybrid Incompatibilities in
19 drosophila. *Genetics* 163(3):955-972.
- 20 38. Gardner K, Buerkle A, Whitton J, & Rieseberg L (2000) Inferring epistasis in
21 wild sunflower hybrid zones. *Epistasis and the evolutionary process.*, eds Wolf
22 JB, Brodie ED, III, & Wade MJ (Oxford University Press, New York), pp 264-
23 279.

- 1 39. Schumer M, *et al.* (2014) High-resolution Mapping Reveals Hundreds of Genetic
2 Incompatibilities in Hybridizing Fish Species. *eLife* 3.
- 3 40. Payseur BA & Hoekstra HE (2005) Signatures of reproductive isolation in
4 patterns of single nucleotide diversity across inbred strains of mice. *Genetics*
5 171(4):1905-1916.
- 6 41. Servedio MR (2000) Reinforcement and the genetics of nonrandom mating.
7 *Evolution* 54(1):21-29.
- 8 42. Servedio MR & Noor MAF (2003) The role of reinforcement in speciation:
9 Theory and data. *Annu Rev Ecol Evol S* 34:339-364.
- 10 43. Kirkpatrick M & Servedio MR (1999) The reinforcement of mating preferences
11 on an island. *Genetics* 151(2):865-884.
- 12 44. Nosil P & Yukilevich R (2008) Mechanisms of reinforcement in natural and
13 simulated polymorphic populations. *Biol J Linnean Soc* 95(2):305-319.
- 14 45. Fisher HS, Mascuch SJ, & Rosenthal GG (2009) Multivariate male traits misalign
15 with multivariate female preferences in the swordtail fish, *Xiphophorus*
16 *birchmanni*. *Anim Behav* 78(2):265-269.
- 17 46. Rosenfield JA & Kodric-Brown A (2003) Sexual selection promotes
18 hybridization between Pecos pupfish, *Cyprinodon pecosensis* and sheepshead
19 minnow, *C-variegatus*. *J Evolution Biol* 16(4):595-606.
- 20 47. Culumber ZW, Ochoa OM, & Rosenthal GG (In press) Assortative mating and
21 the maintenance of population structure in a natural hybrid zone. *Am Nat*.
- 22 48. Ganem G, Lital C, & Lenormand T (2008) Variation in mate preference across a
23 house mouse hybrid zone. *Heredity* 100(6):594-601.

- 1 49. Matute DR, Butler IA, Turissini DA, & Coyne JA (2010) A Test of the Snowball
2 Theory for the Rate of Evolution of Hybrid Incompatibilities. *Science*
3 329(5998):1518-1521.
- 4 50. Corbett-Detig RB, Zhou J, Clark AG, Hartl DL, & Ayroles JF (2013) Genetic
5 incompatibilities are widespread within species. *Nature* 504(7478):135-137.
- 6 51. Harushima Y, Nakagahra M, Yano M, Sasaki T, & Kurata N (2001) A genome-
7 wide survey of reproductive barriers in an intraspecific hybrid. *Genetics*
8 159(2):883-892.
- 9 52. Janousek V, *et al.* (2012) Genome-wide architecture of reproductive isolation in a
10 naturally occurring hybrid zone between *Mus musculus musculus* and *M. m.*
11 *domesticus*. *Mol Ecol* 21(12):3032-3047.
- 12 53. Turner LM, White MA, Tautz D, & Payseur BA (2014) Genomic Networks of
13 Hybrid Sterility. *PLoS Genet* 10(2).
- 14 54. Bomblies K, *et al.* (2007) Autoimmune response as a mechanism for a
15 Dobzhansky-Muller-type incompatibility syndrome in plants. *PLoS Biol*
16 5(9):1962-1972.
- 17 55. Sweigart AL (2010) Simple Y-Autosomal Incompatibilities Cause Hybrid Male
18 Sterility in Reciprocal Crosses Between *Drosophila virilis* and *D. americana*.
19 *Genetics* 184(3):779-U225.
- 20 56. Lee H-Y, *et al.* (2008) Incompatibility of Nuclear and Mitochondrial Genomes
21 Causes Hybrid Sterility between Two Yeast Species. *Cell* 135(6):1065-1073.

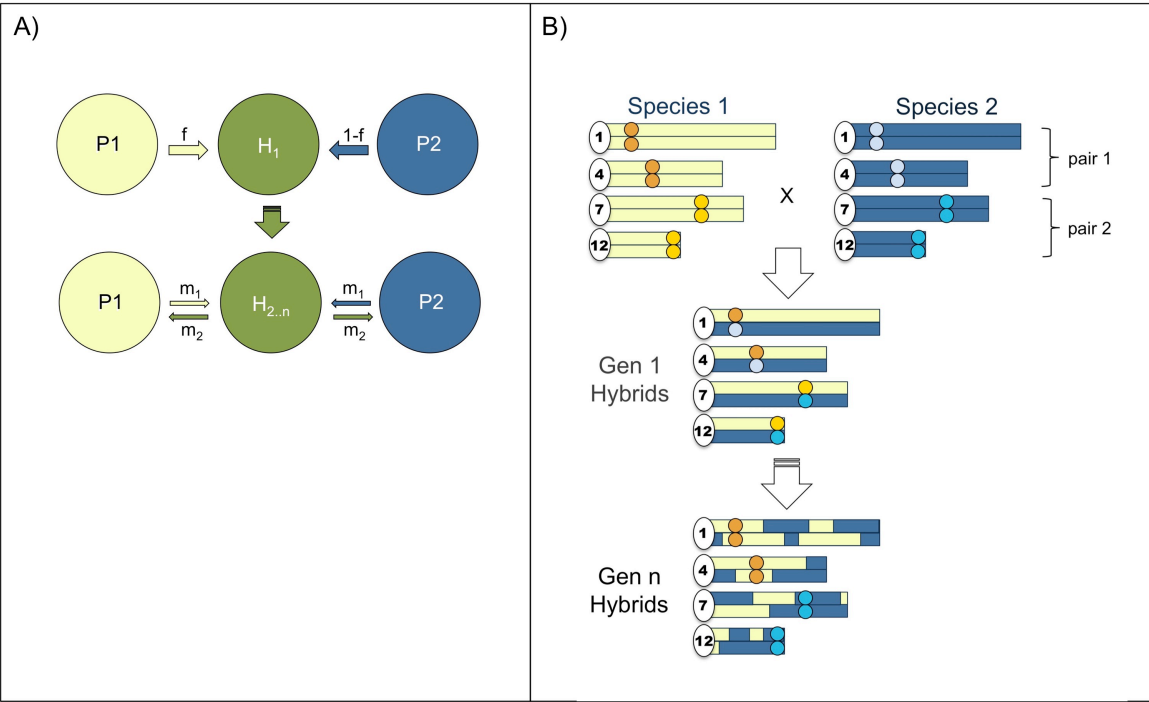
- 1 57. Rawson PD & Burton RS (2002) Functional coadaptation between cytochrome c
2 and cytochrome c oxidase within allopatric populations of a marine copepod.
3 *Proc Natl Acad Sci USA* 99(20):12955-12958.
- 4 58. Johnson NA (2010) Hybrid incompatibility genes: remnants of a genomic
5 battlefield? *Trends Genet* 26(7):317-325.
- 6 59. Crespi B & Nosil P (2013) Conflictual speciation: species formation via genomic
7 conflict. *Trends Ecol Evol* 28(1):48-57.
- 8 60. Barbash DA, Roote J, & Ashburner M (2000) The *Drosophila melanogaster*
9 hybrid male rescue gene causes inviability in male and female species hybrids.
10 *Genetics* 154(4):1747-1771.
- 11 61. Maheshwari S & Barbash DA (2011) The Genetics of Hybrid Incompatibilities.
12 *Ann Rev Genet* 45:331-355.
- 13 62. Staubach F, *et al.* (2012) Genome Patterns of Selection and Introgression of
14 Haplotypes in Natural Populations of the House Mouse (*Mus musculus*). *PLoS*
15 *Genet* 8(8).
- 16 63. Pritchard VL & Edmands S (2013) The genomic trajectory of hybrid swarms:
17 outcomes of repeated crosses between populations of *Tigriopus californicus*.
18 *Evolution* 67(3):774-791.
- 19 64. Trier CN, Hermansen JS, Sætre G-P, & Bailey RI (2014) Evidence for Mito-
20 Nuclear and Sex-Linked Reproductive Barriers between the Hybrid Italian
21 Sparrow and Its Parent Species. *PLoS Genet* 10(1).

- 1 65. Schwander T, Suni SS, Cahan SH, & Keller L (2008) Mechanisms of
2 reproductive isolation between an ant species of hybrid origin and one of its
3 parents. *Evolution* 62(7):1635-1643.
- 4 66. Gross BL & Rieseberg LH (2005) The ecological genetics of homoploid hybrid
5 speciation. *J Hered* 96(3):241-252.

6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23

1 **Figures**

2



3
4 **Fig. 1. Schematic of the simplest hybrid speciation scenario.** The simplest model of
5 hybrid speciation evolves in a hybrid swarm (A) via fixation of parental genetic
6 incompatibility pairs in opposite directions (B). f is the proportion of the hybrid (H)
7 population colonized by parent 1 (P1), $m_{1,2}$ denotes migration rates between the parental
8 and hybrid populations over n generations.

9

10

11

12

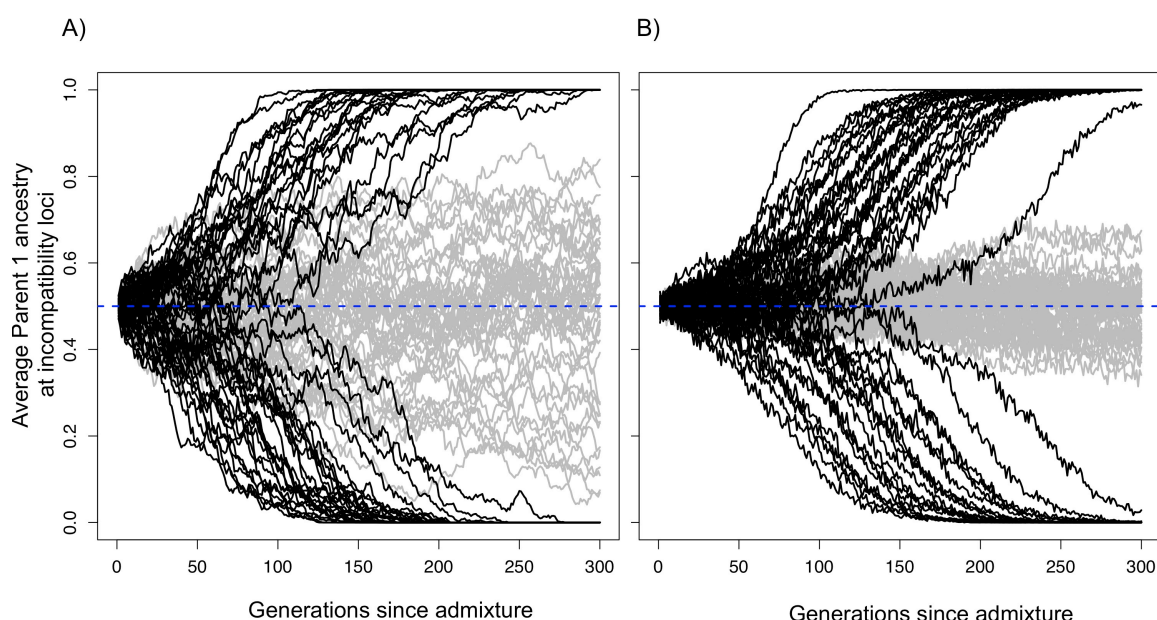


Fig. 2. Hybrid populations rapidly fix for hybrid incompatibility locus pairs.

Selection drives hybrid incompatibility loci to fixation, even when a hybrid population forms at equal admixture proportions ($f=0.5$). Black lines show average parent 1 ancestry at a hybrid incompatibility pair ($h=0.5$, $s=0.1$) in a population of (A) $N=1,000$ or (B) $N=10,000$ diploid individuals. Gray lines show results for this same population size with no selection. Because of this behavior, when two incompatibility pairs exist they may fix for opposite parents, resulting in reproductive isolation of hybrids from both parents (see Fig. 1B).

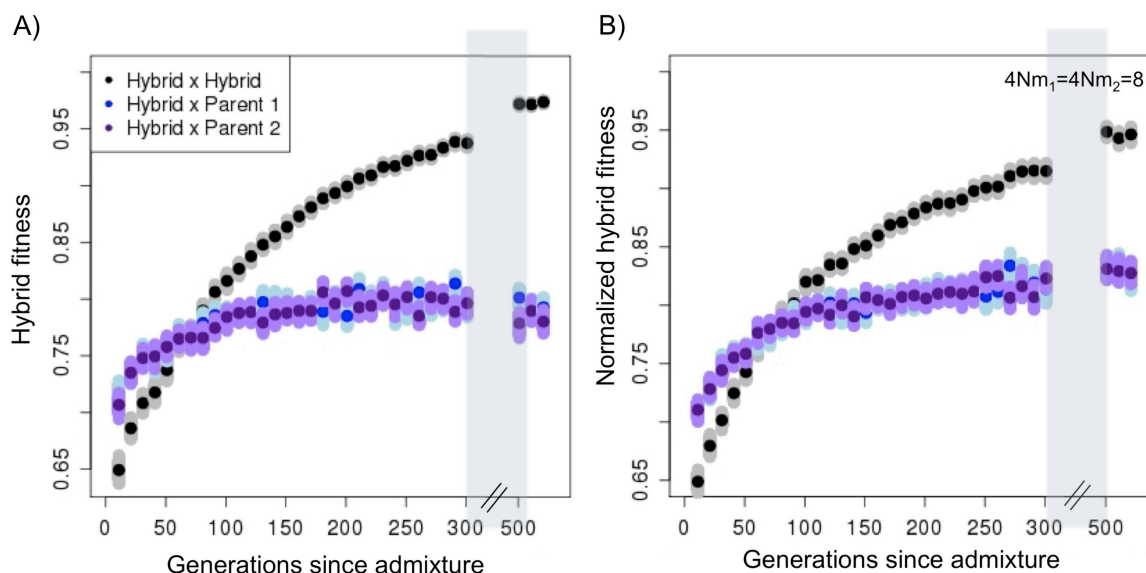


Fig. 3. Hybrid populations rapidly develop reproductive isolation from both parental species, even in the presence of migration. (A) Once hybrid populations diverge in ancestry at hybrid incompatibility loci from parental populations, individual hybrids have higher fitness on average when they mate with other hybrids in their population compared to either parent (B) even with ongoing migration ($4Nm_1=4Nm_2=8$) from parental populations. Dark point indicates mean fitness, smears indicate 1,000 bootstrap resamplings of the mean. Simulation parameters: 100 replicates per time point, $N=1000$, 20 hybrid incompatibility pairs, s_1 , s_2 and h drawn from a distribution (see details in Supporting Information 5D).

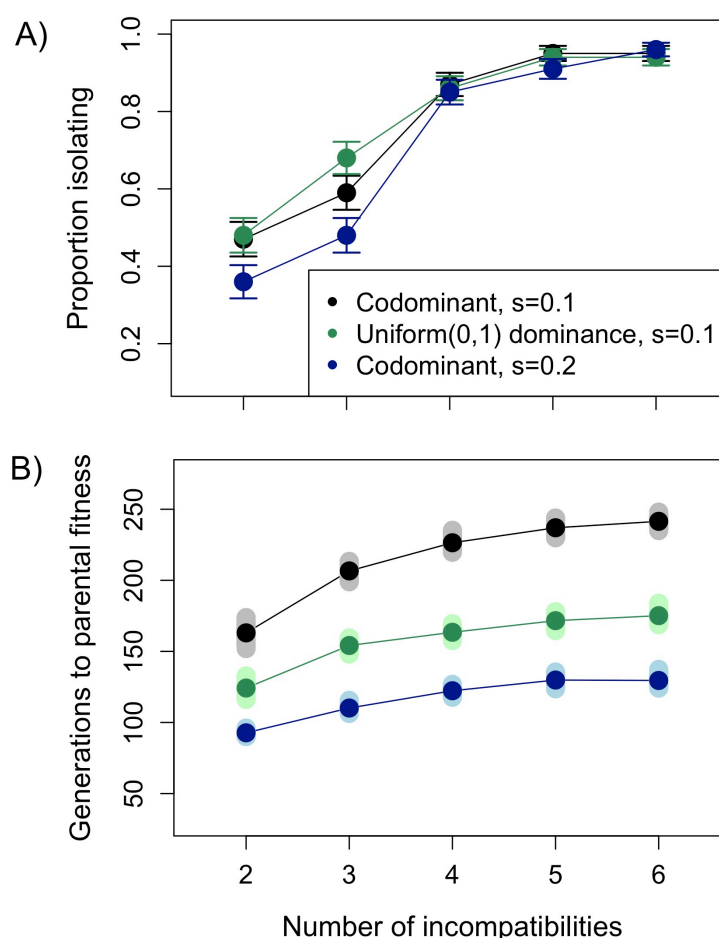


Fig. 4. Relationship between the number of hybrid incompatibility pairs and probability of evolving isolation from both parents. With an increasing number of hybrid incompatibility pairs, reproductive isolation from both parents increases in likelihood (A) but populations require longer periods of time to reach parental fitness levels (B). In these simulations two to six hybrid incompatibility pairs distinguish the hybridizing species ($s=0.1$) and hybrid populations formed at equal admixture proportions ($f=0.5$, 1000 diploid individuals). In (A) error bars show two standard deviations of the mean; in (B) smears show 1000 bootstrap resamplings of the mean. Results based on 500 simulations.