doi:10.1111/evo.14626

On the impermanence of species: The collapse of genetic incompatibilities in hybridizing populations

Tianzhu Xiong[1,2] **and James Mallet**[1]

[1] *Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA*

[2] *E-mail: txiong@g.harvard.edu*

Received February 7, 2022

Accepted July 23, 2022

Species pairs often become genetically incompatible during divergence, which is an important source of reproductive isolation. An idealized picture is often painted where incompatibility alleles accumulate and fix between diverging species. However, recent studies have shown both that incompatibilities can collapse with ongoing hybridization, and that incompatibility loci can be polymorphic within species. This paper suggests some general rules for the behavior of incompatibilities under hybridization. In particular, we argue that redundancy of genetic pathways can strongly affect the dynamics of intrinsic incompatibilities. Since fitness in genetically redundant systems is unaffected by introducing a few foreign alleles, higher redundancy decreases the stability of incompatibilities during hybridization, but also increases tolerance of incompatibility polymorphism within species. We use simulations and theories to show that this principle leads to two types of collapse: in redundant systems, exemplified by classical Dobzhansky–Muller incompatibilities, collapse is continuous and approaches a quasi-neutral polymorphism between broadly sympatric species, often as a result of isolation-by-distance. In nonredundant systems, exemplified by co-evolution among genetic elements, incompatibilities are often stable, but can collapse abruptly with spatial traveling waves. As both types are common, the proposed principle may be useful in understanding the abundance of genetic incompatibilities in natural populations.

KEY WORDS: **Genetic redundancy, hybridization, incompatibility, purging of incompatibilities, reproductive isolation.**

Speciation, the branching process whereby lineages split and persist through time, is often accompanied by the evolution of incompatibilities, or "intrinsic postzygotic isolation" (Coughlan & Matute, 2020; Coyne & Orr, 2004; Dagilis et al., 2019; Orr & Turelli, 2001). Hybrid incompatibility may not be the direct cause of initial divergence, but its presence is considered important for maintaining species integrity when different lineages meet and exchange genes (Coughlan & Matute, 2020). It has been argued that "unlike … pre-mating isolation and extrinsic postzygotic isolation, intrinsic postzygotic isolation is likely to be permanent" (Coyne & Orr, 2004; Muller, 1942). The conceptual picture of genetic incompatibilities has mostly been explored using the classical two-locus incompatibility model (Bateson, 1909; Dobzhansky, 1937; Muller, 1942; Orr & Turelli, 2001), where derived alleles arise on different loci in each species, and these alleles interact negatively to produce hybrids with lower fitness. We here use the term "classical Dobzhansky–Muller systems" as a synonym for this two-locus model, although it has sometimes been used more generally as a synonym of more complex epistatic incompatibilities.

Recent empirical work on genetic incompatibilities has revealed rather more complex genetic architectures. Incompatibilities in hybrids are generally epistatic, and multiple interacting loci suggest a complex basis of hybrid dysfunction (Courret et al., 2019; Meiklejohn et al., 2018; Rosser et al., 2021). Incompatibilities can also be present as polymorphisms within species (Atlan et al., 2003; Cutter, 2012; Corbett-Detig et al., 2013; Larson et al., 2018; Zuellig & Sweigart, 2018b). In some cases, balancing selection has been invoked to explain widespread incompatibility polymorphisms (Seidel et al., 2008), but in most cases, it is unclear whether incompatibility variation has reached equilibrium in space and time. Incompatibility genes might also flow

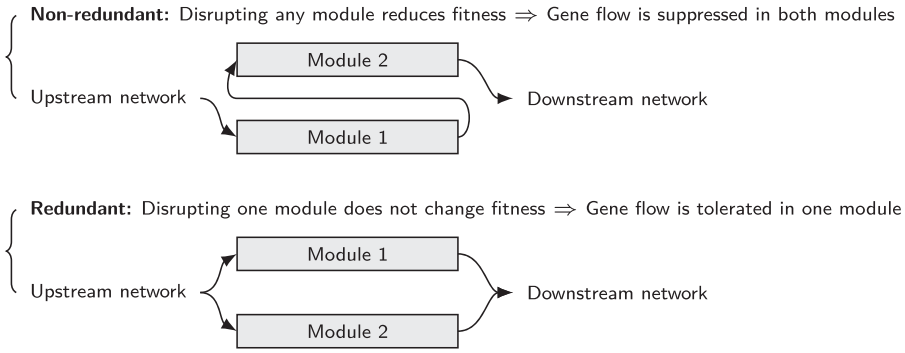

Figure 1. Synopsis of the duality between genetic redundancy and the stability of incompatibilities to hybridization. A "module" refers to a set of factors in a genetic/biochemical network (Schlosser & Wagner, 2004), often associated with an input–output relationship (e.g., a regulatory pathway of downstream genes).

between young species, as has been confirmed or suspected in a number of hybridizing systems (Larson et al., 2018; Presgraves & Meiklejohn, 2021; Rosser et al., 2019). Since hybridization is a frequent phenomenon among recently diverged and incipient species (Mallet, 2005), understanding incompatibilities subject to gene flow is crucial for elucidating how incompatibility evolves in nature.

Over the past few decades, new theories have been proposed for the accumulation of incompatibility alleles between a pair of isolated, panmictic populations (Dagilis et al., 2019; Orr & Turelli, 2001; Schiffman & Ralph, 2022). Mostly, polymorphism is viewed as a transient phenomenon of little importance that occurs in isolated populations. In these "accumulation" theories, hybridization resembles that seen in laboratory crosses, but is assumed to have little impact on parental species. Theoretical analysis of incompatibility dynamics under gene flow has hitherto been limited to a single population receiving migrants from a large, static species, often ignoring stochasticity (Bank et al., 2012; Blanckaert & Hermisson, 2018; Blanckaert et al., 2020). The complexity of incompatibility systems in natural populations is likely much higher, especially if nonequilibrium patterns of hybridization and population structures are considered. Of particular interest are patterns generated by intrinsic incompatibilities (i.e., without ecological, or extrinsic selection), because intrinsic properties are likely to be shared among different taxa having a similar genetic basis of incompatibility. Ecological selection is likely important, but can be added as an extra layer when necessary. If such intrinsic patterns evolve relatively rapidly, they will account for a nontrivial component of observed incompatibilities in nature, as empirical evidence tends to show (Coyne & Orr, 2004; Coughlan & Matute, 2020).

Many incompatibilities are polygenic and epistatic, which obstructs the formulation of simple, robust principles that apply independently of model details. Therefore, we first present our intuition behind general models underlying a theoretical analysis: that patterns of incompatibilities under hybridization can be broadly classified by the level of genetic redundancy (Fig. 1). Then we demonstrate qualitative differences of incompatibility collapse by analyzing and simulating each class of incompatibility systems under increasingly complex population structures.

Theory

INTUITION: THE DUALITY BETWEEN GENETIC REDUNDANCY AND THE STABILITY OF INCOMPATIBILITIES TO HYBRIDIZATION

Genetically redundant incompatibilities

Redundant components such as duplicated genes or pathways greatly facilitate the evolution of genetic incompatibilities, as they produce extra degrees of freedom even for species under stabilizing selection (Haag, 2007; Lynch & Force, 2000). Classical Dobzhansky–Muller incompatibilities represent the simplest form of redundant system, because substituting a single derived allele into the ancestral background has no fitness cost: the fitness of the ancestral genotype is robust to small-scale disruptions (Fig. 1, bottom). However, this redundancy also de-stabilizes incompatibilities in the face of gene flow, first shown for classical Dobzhansky–Muller systems, where continuing hybridization can lead to their total collapse (Barton & Bengtsson, 1986). Empirical examples of redundant incompatibilities are not rare. In *Mimulus*, duplicated genes with reciprocal pseudogenization in each species create pseudogene–pseudogene incompatibilities (Zuellig & Sweigart, 2018a), a special case of the Dobzhansky–Muller model that will typically be unstable to hybridization (Bank et al., 2012). In *Arabidopsis thaliana*, the existence of additional gene copies as rescuers of a known pseudogene–pseudogene incompatibility further increases redundancy and two pseudogenes can co-exist at arbitrary frequencies within a population (Jiao et al., 2021). In other populations of *A. thaliana*, first-generation hybrid necrosis was mapped to a

large array of dominantly interacting immune-response genes, which arise via multiple cycles of gene duplication and gene loss (Chae et al., 2014). In shrews, lineages of *Sorex araneus* with different Robertsonian chromosome fusions, a karyotypic example of Dobzhansky–Muller incompatibility, collapse into fully acrocentric, globally compatible chromosomes within the hybrid zone (Hatfield et al., 1992). Genetic redundancy also frequently appears in both *trans*- and *cis*-regulation that may be important for the development of incompatibilities (True & Haag, 2001). For instance, distinct transcription factors could respond to similar signals, and control a similar set of target genes (AkhavanAghdam et al., 2016; Hu et al., 2007), and some enhancers are also functionally redundant (Barolo, 2012; True & Haag, 2001). In principle, if a biological function is redundantly controlled by several independent genetic modules, it may be robust to disruptions of a subset of them (Fig. 1, bottom). Gene flow can thus occur in at least a subset of modules, which destabilizes incompatibilities.

Genetically nonredundant incompatibilities

Conversely, incompatibilities lacking genetic redundancy are typically more resistant to hybridization (Fig. 1, top), and are bistable. A classic example is the co-evolution between *cis*- and *trans*- acting regulatory elements, where any recombined genotype suffers from mis-regulation (Barriere et al., 2012; Mack & Nachman, 2017; Prager & Wilson, 1975; Porter & Johnson, 2002). Since intermediate genotypes are suppressed, incompatibilities are stable in the presence of gene flow (Bazykin, 1973; Barton, 1979; Lindtke & Buerkle, 2015). This argument could be extended to other examples, such as gene-dosage imbalance (Josefsson et al., 2006), cyto-nuclear interactions (Chou & Leu, 2010), selfish genetic elements with their suppressors (Case et al., 2016; Phadnis & Orr, 2009), or reciprocal sign epistasis (Bazykin, 1973; Poelwijk et al., 2011). Interestingly, some incompatibilities might be partially redundant. For instance, polygenic threshold incompatibilities are frequently observed in *Drosophila* (Liénard et al., 2016; Morán & Fontdevila, 2014), where strong hybrid dysfunction occurs only when the amount of introgressed DNA exceeds a certain fraction. These systems might tolerate a moderate level of segregation involving incompatibility alleles (Morimoto et al., 2020).

We emphasize that duplicated components in genomes may vary over time (Cotton & Page, 2005; Moore & Purugganan, 2005), as does genetic redundancy. Thus, incompatibility factors evolved by redundant processes might lack redundancy in some hybrid populations. Consequently, the proposed conceptual duality should be between the "current" level of genetic redundancy and the stability of genetic incompatibilities to hybridization. This duality forms the basis of our models below.

Table 1. Glossary of major symbols.

Symbol	Meaning
G	The state space of genotype distributions
$\mathbf{g}$	A particular genotype distribution (an element in G)
$\boldsymbol{\mu}$	Mortality rates of all genotypes
s	The increase in mortality rate due to incompatibilities
α	Mating rate
m	The probability of hybridization between sympatric species (assumed to be small)
m'	The probability of nearest-neighbor migration in discrete populations
$\mathbf{Q}$	The stoichiometric matrix mapping the distribution of matings to offspring genotypes
r	Recombination probability in the two-locus model
c_{allele}	The count of a particular allele in a diploid genotype (c_A, c_B, etc. Takes value in $\{0, 1, 2\}$)
p	The frequency of a particular allele
p_s	The probability that a random sample never hit regions of sympatry throughout its genealogy
σ^2	The variance of a one-dimensional dispersal kernel ($\sigma^2 = m'$ in discrete populations)
L	The size of a linear habitat
$\mathbf{x}, t$	Coordinates in space and time
N	Diploid population size

A TOPOLOGICAL DEFINITION OF REDUNDANCY

We start by considering two genetically incompatible species, 1 and 2. Let Ω be the space of all possible genotypes that can be created by recombination and union of all haplotypes that preexist in either species 1 or species 2. The set of all probability distributions on Ω is denoted by G, which corresponds to the general state space of genotype distributions. Elements in G are denoted by $\mathbf{g}$. We use $\mathbf{g}_1$ and $\mathbf{g}_2$ as the genotype distributions in the two species, respectively. We assume that generations are overlapping, and microscopic changes are governed by an individual-based birth-death process (see Supporting Information). A glossary of symbols and variables can be found in Table 1). If genetic drift is neglected, evolution in a single randomly mating population is driven by viability selection and reproduction. Its dynamics in the genotypic space G can be written concisely as:

$$\partial_t \mathbf{g} = \underbrace{[\bar{\mu}_{\mathbf{g}}\, \mathbf{I} - \mathrm{diag}(\boldsymbol{\mu})]\, \mathbf{g}}_{\text{Selection by incompatibility}} + \underbrace{\alpha\, \mathbf{Q}\, (\mathbf{g} \otimes \mathbf{g})}_{\text{Reproduction}}, \tag{1}$$

where α is the rate of mating, $\mathbf{Q}$ is the stoichiometric matrix relating the distribution of offspring genotypes to the distribution of matings (given by the tensor product $\mathbf{g} \otimes \mathbf{g}$), $\mathbf{I}$ is the identity matrix, vector $\boldsymbol{\mu}$ contains the mortality rates of all genotypes, and its average value over a genotype distribution $\mathbf{g}$ is denoted as $\bar{\mu}_{\mathbf{g}}$.

As a first approximation, we assume a "holey landscape" fitness model (Gavrilets, 1997). Ecological selection is not considered as we are here primarily interested in intrinsic properties of incompatibilities. Under the holey landscape model, mortality rates of genotypes are either normal or very high. To model the influence of hybridization from species 2 on species 1, split the reproduction term to include a small probability m of interspecific hybridization, and we have the following equation governing the genotype distribution $\mathbf{g}_1$:

$$\partial_t \mathbf{g}_1 = \underbrace{\left[\bar{\mu}_{\mathbf{g}_1}\,\mathbf{I} - \mathrm{diag}(\boldsymbol{\mu})\right]\mathbf{g}_1}_{\text{Selection by incompatibility: } \mathbf{f}_s(\mathbf{g}_1)} + \underbrace{\alpha(1-m)\,\mathbf{Q}\,(\mathbf{g}_1 \otimes \mathbf{g}_1)}_{\text{Within-species reproduction: } \mathbf{f}_w(\mathbf{g}_1)} + \underbrace{\alpha m\,\mathbf{Q}\,(\mathbf{g}_2 \otimes \mathbf{g}_1)}_{\text{Hybridization: } \mathbf{f}_h(\mathbf{g}_1,\mathbf{g}_2)} \quad (2)$$

Note that the probability of interspecific gene flow under this formulation is $m/2$, because each first-generation hybrid only has half of its DNA inherited from its foreign parent. Consider the following set of fixed points:

$$\widehat{\mathsf{G}} = \{\mathbf{g} | \partial_t \mathbf{g} = \mathbf{f}_s(\mathbf{g}) + \mathbf{f}_w(\mathbf{g})\} \quad (3)$$

Since multiple genotypes are equally fit under the holey landscape model, this set might become degenerate (i.e., not composed entirely of isolated points). If parental genotype distributions $\mathbf{g}_1$ and $\mathbf{g}_2$ are path-connected in $\widehat{\mathsf{G}}$, then it is possible to perturb $\mathbf{g}_1$ along such paths to $\mathbf{g}_2$ without leaving the equilibrium, assuming weak hybridization. The existence of a path-connected $\widehat{\mathsf{G}}$ is not always guaranteed, but its presence is what we mean by redundant incompatibilities. There are good reasons for such definition: Genetic redundancy implies that maintenance of high fitness does not require all components of the system to work (Nowak et al., 1997). This adds some free dimensions to the dynamics that are unconstrained by fitness. The existence of a path-connected $\widehat{\mathsf{G}}$ is the consequence of such free dimensions. An extreme version of redundancy is neutral evolution, which exists if the entire state space becomes $\widehat{\mathsf{G}}$. In contrast, when $\widehat{\mathsf{G}}$ consists only of isolated points, the evolutionary dynamics switches to jumps similar to the shifting-balance process (Wright, 1932). Thus, there are two different regimes of dynamics depending on whether genetic redundancy exists (Figure 2a):

Dynamics of nonredundant incompatibilities

If $\widehat{\mathsf{G}}$ is not path-connected, under the holey landscape model, we can asymptotically eliminate only the time-scale of selection by incompatibility. Such systems are governed by:

$$\partial_t \mathbf{g}_1 \approx \lim_{\tau\to+\infty} e^{\tau \mathbf{J}_{\mathbf{f}_s(\mathbf{g}_1)}}\left[\mathbf{f}_w(\mathbf{g}_1) + \mathbf{f}_h(\mathbf{g}_1,\mathbf{g}_2)\right] := \mathbf{f}_{\mathrm{non}}(\mathbf{g}_1,\mathbf{g}_2), \quad (4)$$

where $\mathbf{J}$ stands for the Jacobian matrix at equilibrium points of the dynamical forces specified by its subscript (see Supporting Information). Such incompatibilities are resolved by an abrupt jump (i.e., a genetic peak-shift. A two-locus example is given in Figure 2b).

Dynamics of redundant incompatibilities

If $\widehat{\mathsf{G}}$ is path-connected, under suitable conditions we can asymptotically eliminate time-scales of selection and of within-species reproduction, and the system evolves under the time-scale of hybridization:

$$\partial_t \mathbf{g}_1 \approx \lim_{\tau\to+\infty} e^{\tau \mathbf{J}_{\mathbf{f}_s(\mathbf{g}_1)+\mathbf{f}_w(\mathbf{g}_1)}}\left[\mathbf{f}_h(\mathbf{g}_1,\mathbf{g}_2)\right] := \mathbf{f}_{\mathrm{red}}(\mathbf{g}_1,\mathbf{g}_2) \quad (5)$$

Such incompatibilities will collapse by incremental perturbations due to hybridization (see Figure 2b for a two-locus example).

Stochasticity

In finite populations, Equations (4) and (5) are also affected by genetic drift in the birth-death process. The influence of stochasticity is essential to understanding many aspects of collapse dynamics, and will be treated analytically in two locus incompatibility models presented below. For more complicated systems, we show simulated results.

SPECIFIC MODELS

Inspired by realistic molecular mechanisms, we here define four specific incompatibility models with different levels of redundancy (Table 2), on which we base subsequent results. Simpler models (I and II) are mainly used for the analytical treatment of collapse dynamics, which have already been studied somewhat (Lindtke & Buerkle, 2015). More complex models (III and IV) conceived for highly redundant incompatibilities are studied here using simulations in SLiM-3.6 (Haller & Messer, 2019) to demonstrate the robustness of certain quasi-neutral behaviors in spatially extended populations. At most four bi-allelic loci in diploid species will be considered, and we track alleles that are represented by capitalized letters (A, B, C, D). Under the birth-death formulation, these models have the following form of mortality rates:

$$\mu = \mu_0 + s\mathbb{I}(\text{Incompatibility condition}), \quad (6)$$

where $\mathbb{I}$ is the indicator function, s is the increase in mortality rates by incompatibility. Further, let c_{allele} be the count of the designated allele in a diploid genotype.

Model I: Two-locus bistable systems (non-redundant)

Bistable systems with two loci could represent co-evolution between regulatory elements in a pathway under stabilizing selection. In the model, we assume fitness reduction is due to additive

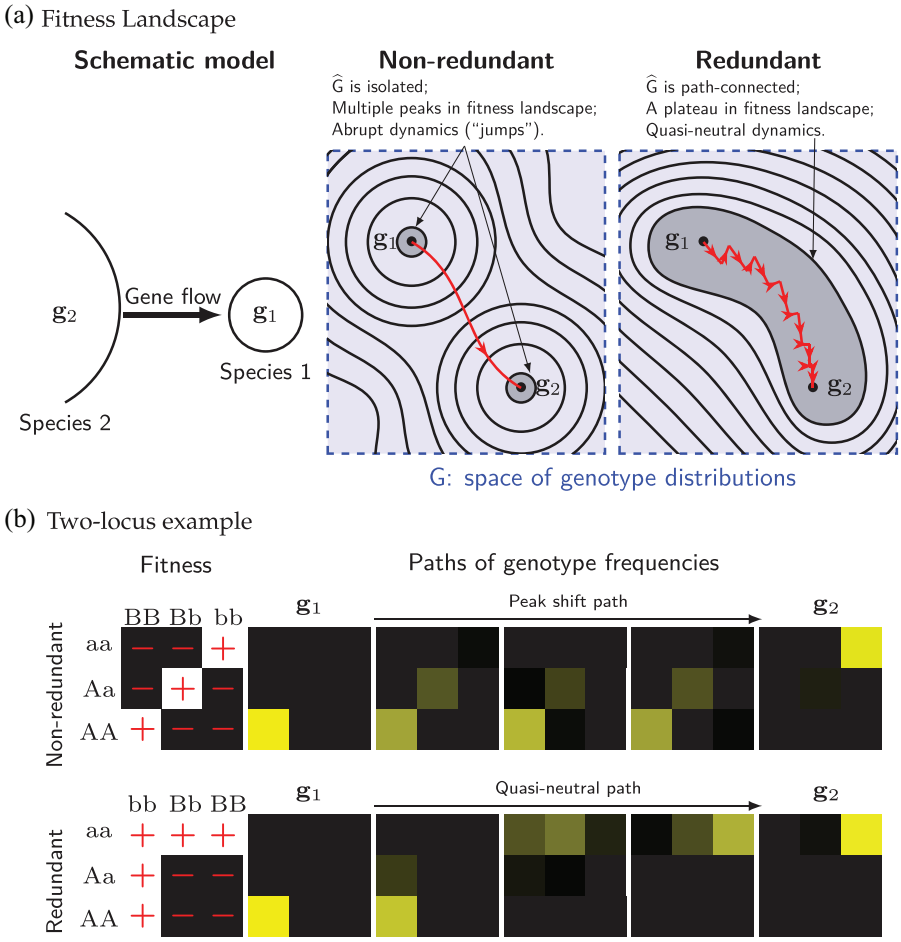

Figure 2. **Schematic differences between the collapse dynamics of redundant and non-redundant incompatibilities. For simplicity, species 1 is subject to unidirectional gene flow from species 2. (A) Representation of dynamics in the general model. Red lines depict the trajectories of genotype distributions in species 1 subject to unidirectional gene flow from species 2. Contours depict the fitness landscape in the genotype distribution space G. For redundant incompatibilities, collapse trajectories form a quasi-neutral path on the fitness plateau. For nonredundant incompatibilities, collapse requires a stochastic jump. (B) Explicit examples of dynamics in two-locus models (see Section "Specific models". Nonredundant: Model I; Redundant: Model II).**

dosage-imbalance between two loci, which leads to the incompatibility criterion:

$$c_A \neq c_B \tag{7}$$

This bistable system broadly represents behaviors of nonredundant incompatibilities that are intrinsically symmetric, because mortality rates are invariant when exchanging labels "A" and "B".

Model II: Two-locus unistable Dobzhansky–Muller systems (redundant)

The classical Dobzhansky–Muller system involves two derived alleles (A and B) on two loci, each behaving neutrally on its own. They interact negatively when combined in hybrids. For simplicity, negative epistasis is assumed to be fully dominant, so that the incompatibility criterion becomes:

$$c_A c_B > 0 \tag{8}$$

Model III: Highly redundant Dobzhansky–Muller systems

This redundant system involves duplicating the same unistable locus. For example, consider four pseudogenized alleles (A, B, C, D) of a duplicated gene, where each gene copy also has functional homologs (a, b, c, d). Each gene is unistable on its own because its functional allele will always replace the pseudogenized one under selection. Mortality rate is normal whenever a

Table 2. Specific models of incompatibilities.

Model	Initial diploid genotypes		Incompatibility condition	Recombination	Redundancy
	Species 1	Species 2			
I	AABB	aabb	$c_A \neq c_B$	r	Non-redundant
II	AAbb	aaBB	$c_A c_B > 0$	r	Redundant
III	AABBccdd	aabbCCDD	$c_A = c_B = c_C = c_D = 2$	Unlinked	Highly redundant
IV	AABBCCDD	aabbccdd	$\lvert(c_A - 1)(c_B - 1)\rvert + \lvert(c_C - 1)(c_D - 1)\rvert = 0$	Unlinked	Highly redundant

single functional copy is present in an individual, and the incompatibility criterion is:

$$c_A = c_B = c_C = c_D = 2 \tag{9}$$

This model can also be seen as an extension of the two-locus Dobzhansky–Muller holey landscape model with recessive interactions.

Model IV: Highly redundant multistable systems

This system considers a type of genetic redundancy in pathway duplication where each divergent pathway produces multi-stable incompatibility and is non-redundant, but the joint incompatibility of all such pathways becomes redundant since multiple pathways serve the same function. In the main text, we model two pathways with identical functions, each pathway contains two bi-allelic loci (first pathway: A/a & B/b, second pathway: C/c & D/d), which disrupt the pathway's function whenever a locus becomes heterozygous. The criterion for incompatibility is:

$$\lvert(c_A - 1)(c_B - 1)\rvert + \lvert(c_C - 1)(c_D - 1)\rvert = 0 \tag{10}$$

Results

THE COLLAPSE OF INCOMPATIBILITIES IN A SINGLE POPULATION

Consider a population of species 1 experiencing unidirectional gene flow from species 2. Time courses to the collapse of incompatibilities defined by Models I and II are shown in Figure 3. The average time to collapse $\langle t_c \rangle$ is defined as the first time that every individual in the peripheral lineage becomes compatible with the external species. Since hybridization is assumed to be a weak effect, it has little effect on population size, which fluctuates near N, the equilibrium value. For simplicity, let $m^* = Nm$ and $r^* = Nr$ be the population rates of migration and recombination following classical population genetic notions.

In the nonredundant Model I, the average time to collapse is approximately:

$$\langle t_c \rangle_{\text{non}} \approx \frac{2\pi}{\alpha m} \exp\left[\frac{1}{2}r^* + \left(m^* + \frac{1}{2}\right)\ln\frac{m}{r}\right] \tag{11}$$

$\langle t_c \rangle_{\text{non}}$ approaches infinity when stochasticity goes to zero (i.e., $N \to \infty$) (see Supporting Information). Here, the effect of stochasticity depends on the population migration rate m^* as well as the population recombination rate r^*.

For the redundant Model II, the average time to collapse is given by:

$$\langle t_c \rangle_{\text{red}} \approx \frac{2m^*}{\alpha m} \frac{\psi(m^* + 1) - \psi(m^* \frac{\alpha}{s} r)}{1 + m^*(1 - \frac{\alpha}{s} r)}, \tag{12}$$

where $\psi(\cdot)$ represents the digamma function. $\langle t_c \rangle_{\text{red}}$ is bounded as a function of N (see Supporting Information). In addition, the effect of stochasticity is mediated only through population migration rate m^*.

Although a single population represents only the simplest population structure, it is sufficient to illustrate the different dynamics of collapse between the two classes of incompatibilities. Redundant incompatibilities collapse continuously (bounded $\langle t_c \rangle$), while nonredundant incompatibilities only collapse abruptly in a peak shift (Wright, 1932) (exponential $\langle t_c \rangle$). This dichotomy extends further when we consider collapse in spatially extended populations (see below).

There are other contrasts between the models. In Model I, first-generation hybrids are normal, while backcross generations tend to be unfit. In Model II, fitness reduction occurs immediately in first-generation hybrids, but dissipates in backcross generations. Consequently, stochasticity (genetic drift) affects the persistence of incompatibilities in opposite ways. In Model I, incompatibility alleles are deterministically stable in the focal population when hybridization is weak. Therefore, weaker stochasticity will strengthen the performance of the deterministic barrier. In Model II, however, stochasticity creates a stronger barrier to collapse, because of the increased likelihood of losing the rare early generation hybrids before their offspring spread.

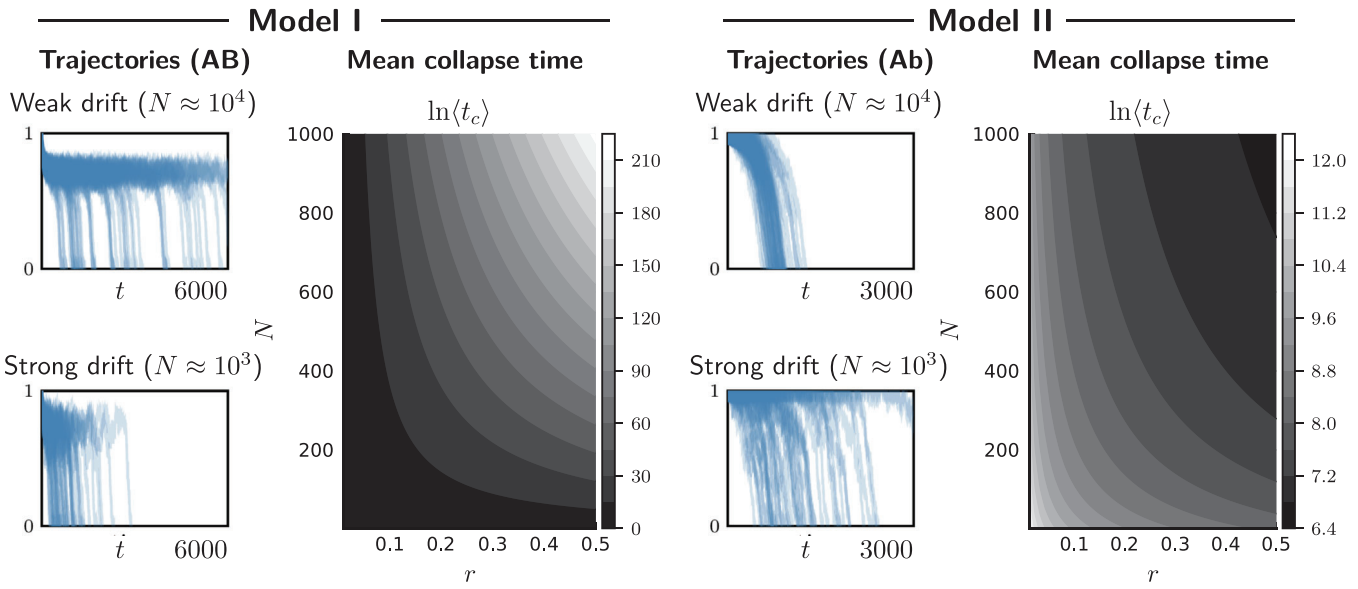

Figure 3. **The collapse of two locus incompatibilities in a population subject to gene flow from a different species. Blue trajectories are frequencies of the native haplotype in the focal population (AB for Model I, and Ab for Model II). The trajectories show results of 50 repeated simulations of the birth–death process. Simulations have the following recombination probabilities: $r = 0.055$ for Model I, and $r = 0.25$ for Model II. Recombination probabilities were chosen such that population size in the range $10^3 \leq N \leq 10^4$ has a non-negligible impact on collapse dynamics. Numerically calculated mean collapse times $\langle t_c \rangle$ using Equations (12) and (11) are shown as heatmaps. All data in this figure correspond to fixed parameters: $s = 10$, $\alpha = 1$, $m = 0.01$.**

When stochasticity is weak, it is increasingly likely to have early generation hybrids that survive their incompatibility to generate backcrosses, which can lead to incompatibility collapse. These effects produce virtually opposite patterns of $\langle t_c \rangle$ as a function of N and r (heatmaps in Figure 3).

THE COLLAPSE OF NONREDUNDANT INCOMPATIBILITIES IN SPACE AND TIME

Nonredundant incompatibilities collapse via stochastic traveling waves between broadly sympatric species

To understand the sympatric collapse of non-redundant incompatibilities, consider a coarse-grained spatial system, where two layers of demes represent local populations of two species (Figure 4a). Each deme is in one of the two discrete states: $+1$ and -1, representing the two parental genotype distributions $\mathbf{g}_1$ and $\mathbf{g}_2$. With nearest-neighbor migration within species (m'), and local hybridization between species (m), each deme can flip its state with a fixed rate depending on the states of its immediate neighbors. A contact probability can be defined for each deme as the fraction of individuals exchanged with demes of opposite incompatibility state, which captures the net influence of other genotypes. For each deme, its rate of flipping increases with a higher contact probability. This simplified model is a reasonable approximation to the full system, as long as the full system is bistable and if the collapse occurs abruptly at a fixed rate. We define a stochastic traveling wave in the discrete system as a biased movement of the boundary between two states. Since such movement requires stochastic flipping of discrete states, wave fronts propagate stochastically in a manner similar to a biased random walk. Due to the requirement for bias, the spatial boundary between opposite states needs to be asymmetric for traveling waves to occur.

In general, the first event to take place is a local collapse (assimilation of discrete states) between a pair of sympatric demes. Following this event, there are two sources of asymmetry in the system which could contribute to the formation of traveling waves. First, if non-redundant incompatibilities are intrinsically asymmetric (i.e., states $+1$ and -1 are not exchangeable without altering the dynamics), traveling waves can form due to gene flow within each species (Barton & Turelli, 2011; Geldhauser & Kuehn, 2020). Second, continuing hybridization creates asymmetry that favors the wave to propagate in the same direction as the initial collapse. The second source of asymmetry means that even a completely symmetric bistable incompatibility can produce traveling waves, which is impossible in the absence of hybridization.

To demonstrate this process, let us analyze the coarse-grained system of Model I and see how traveling waves originate. In this symmetric model, there are three steps to initiate a traveling wave (Figure 4a): In the first step, the system waits for an initial local collapse (panel "Initial state"), which could

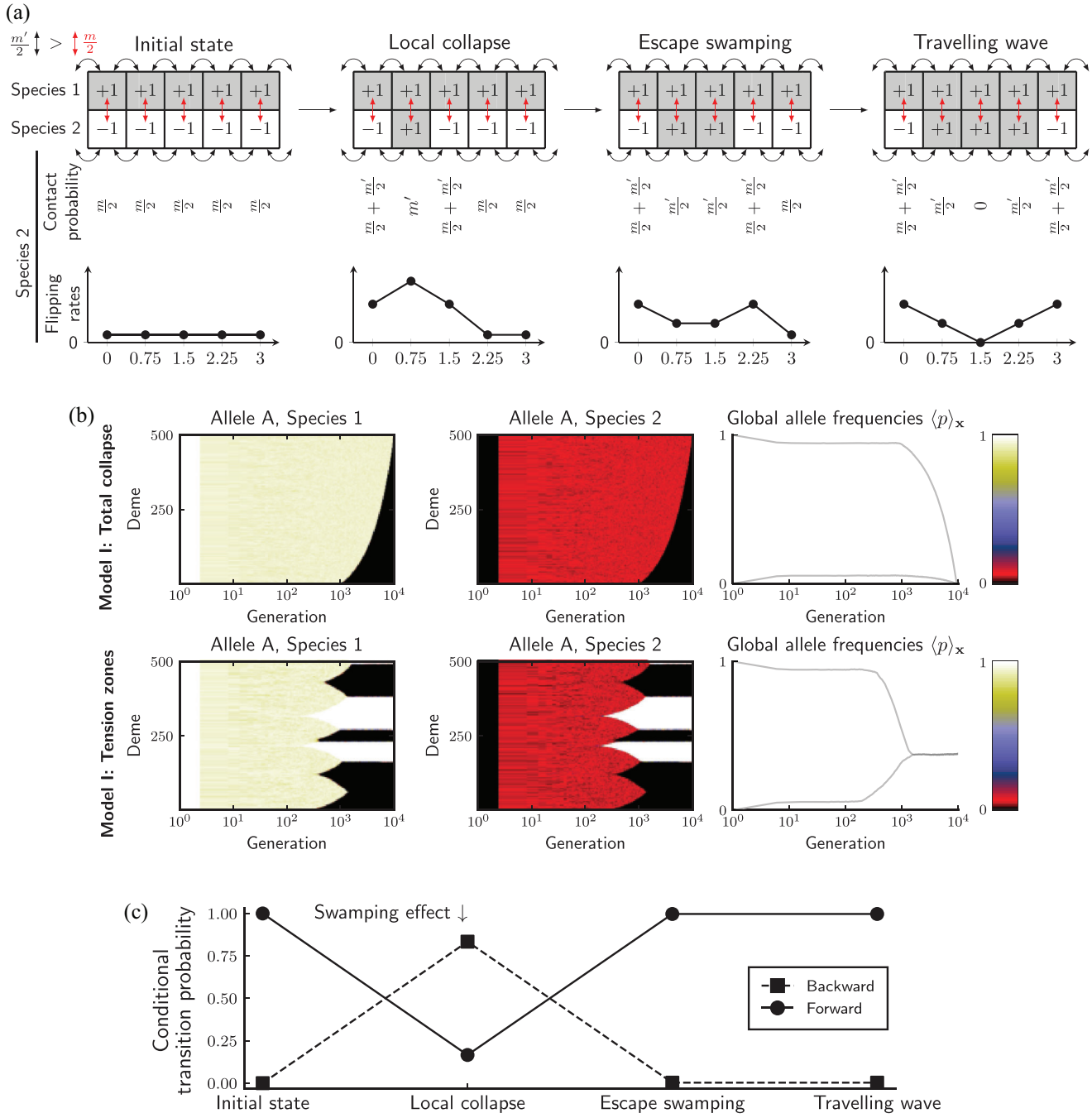

Figure 4. Nonredundant incompatibility collapses via traveling waves in spatial systems. (A) The coarse-grained system of Model I between sympatric species. Once the initial state is broken by a local collapse, the collapsed deme has to survive swamping. If it escapes swamping, the rate of flipping is higher towards the exterior of the collapsed region, which drives the propagation of traveling waves. Gray (+1) and white (-1) boxes correspond to two coarse-grained states of a single deme. m and m' are hybridization and migration probabilities, respectively. (B) Two examples of SLiM simulations of Model I showing that symmetric incompatibilities could collapse completely or form multiple tension zones. Frequencies of allele A were taken every five generations (i.e., $\Delta t = 5$ in Equation 13). Hybridization probability $m = 0.05$, and nearest-neighbor migration probability $m' = 0.1$. Population size in each deme fluctuates around 539 for the iteration where total collapse occurs (top panels), and it fluctuates around 269 where multiple collapses occur (bottom panels). See Supporting Information Figure S3 for the behaviors of all loci. (C) Numerical solutions of conditional transition probabilities among the four states of the entire system. Parameters are the same as SLiM simulations with population sizes fixed at 269. Probabilities are derived from transition rates by inversing $\langle t_c \rangle_{non}$ from Equation S26.

occur in multiple demes across the entire geographic distribution, and different demes might collapse in opposite directions. In the second step, the collapsed region needs to expand spatially to escape swamping. The swamping effect is prominent as a single collapsed deme flanked by two demes of opposite states has the highest rate of flipping back to its original state (panel "Local collapse"). If swamping is prevented, the system proceeds to step three, where traveling waves begin to emerge. This is because demes outside the collapsed region now have the highest rate of flipping states (panel "Escape swamping"). Finally, emerged traveling waves propagate (panel "Traveling wave"). We simulated Model I in a one-dimensional habitat with bidirectional hybridization using SLiM (see Supporting Information). The results (Figure 4b) demonstrate the existence of traveling waves, and the final state of the entire system varies from a total collapse to the formation of multiple tension zones. The likelihood of these outcomes clearly depends on the number of initiated waves, as well as the direction of collapse at each wave front, both of which depend on details of incompatibilities and the level of stochasticity in each deme. The swamping effect can be numerically verified by calculating conditional transition probabilities among the four states under Model I (Figure 4c). These calculations show that the state "local collapse" has a high probability of reversing to the initial condition. However, once swamping is prevented, the system mainly transitions forward to a traveling wave. The presence of traveling waves alters the time-scale of persistence of the incompatibility. On the one hand, if the initial collapse is slow, but traveling waves form readily, a total collapse between the two species is likely to occur. On the other hand, if initial collapse is frequent and occurs in different places, and the incompatibility is symmetric, long-term tension zones can form which separate regions fixed for different incompatibility genotypes.

If two species are only partially sympatric, waves of collapse will be relieved of interspecific hybridization once wave fronts leave sympatric regions. Such fronts might propagate beyond sympatric regions if incompatibilities are asymmetric, and these traveling waves in a single species have been studied extensively in existing literature (Barton, 1979; Barton & Turelli, 2011; Geldhauser & Kuehn, 2020).

THE COLLAPSE OF REDUNDANT INCOMPATIBILITIES IN SPACE AND TIME

Redundant incompatibilities collapse into quasi-neutral spatial polymorphism between broadly sympatric species

At a local spatial scale, the collapse of redundant incompatibilities can be viewed as a situation where two parental genotype distributions converge onto a common point along some quasi-neutral path in the state space G. The final position of this point in the genotypic state space is not unique, because stochasticity can alter the trajectories during the collapse. Consequently, independent realizations of the same collapse will have variable outcomes.

In spatially extended populations, if the habitat is large, genetic drift is decoupled between very distant populations (isolation by distance) (Barton et al., 2013; Wright, 1943). As a result, the immediate outcome of collapse will also differ between distant populations. However, gene flow tightly couples genotypic distributions between adjacent demes, and adjacent populations will evolve to become compatible. Consequently, after collapse, gradual quasi-neutral clines may form in space with no local tension zone.

A signature of quasi-neutrality is that allele-frequency changes will be dominated by genetic drift. Quantitatively, across a small time interval (Δt), the mean-squared allele-frequency-change across space in each species $\langle|\Delta p|^2\rangle_{\mathbf{x}}$ will follow:

$$\langle|\Delta p|^2\rangle_{\mathbf{x}} \approx \left\langle \frac{p(1-p)}{2N}\Delta t \right\rangle_{\mathbf{x}} \tag{13}$$

where $\langle\cdot\rangle_{\mathbf{x}}$ represents averaging across all local populations of that species.

To demonstrate quasi-neutrality, we simulated redundant Models II, III, and IV in a one-dimensional habitat with bidirectional hybridization throughout 500 demes. This is the same spatial setup with gene flow within and between species already tested for Model I (see Figure 4a). Models II–IV have a phase of rapid collapse following the onset of hybridization, after which sympatric local populations share similar allele frequencies between different species (Figure 5). However, in the long-term, allele frequencies still vary substantially across space, and overall allele frequencies in the entire species change only very slowly. When comparing the mean-squared response ($\langle|\Delta p|^2\rangle_{\mathbf{x}}$) against neutral genetic drift ($\Delta t = 5$ generations), we found that the magnitude of allele frequency change deviates from neutrality in the short-term, but fluctuates around neutrality in the long-term, confirming that genetic drift has dominated allele frequency changes after the initial collapse. This marks the existence of a long-term quasi-neutral phase. As all alleles involved in incompatibilities are purged at very low rates in the quasi-neutral phase, it is possible for conspecifics from different local populations to become incompatible. Furthermore, different inbred lines generated from a single local populations could also yield variable levels of incompatibility when crossed with another species, because local polymorphism is tolerated, and is not subject to strong selection.

Another way to interpret quasi-neutrality is to follow a spatial allele frequency cline. In the redundant multistable system (Model IV), despite a lack of tension zones, alleles from different pathways still segregate into broad spatial "bubbles", and fix in an alternating order in space (see banded pattern of frequency after

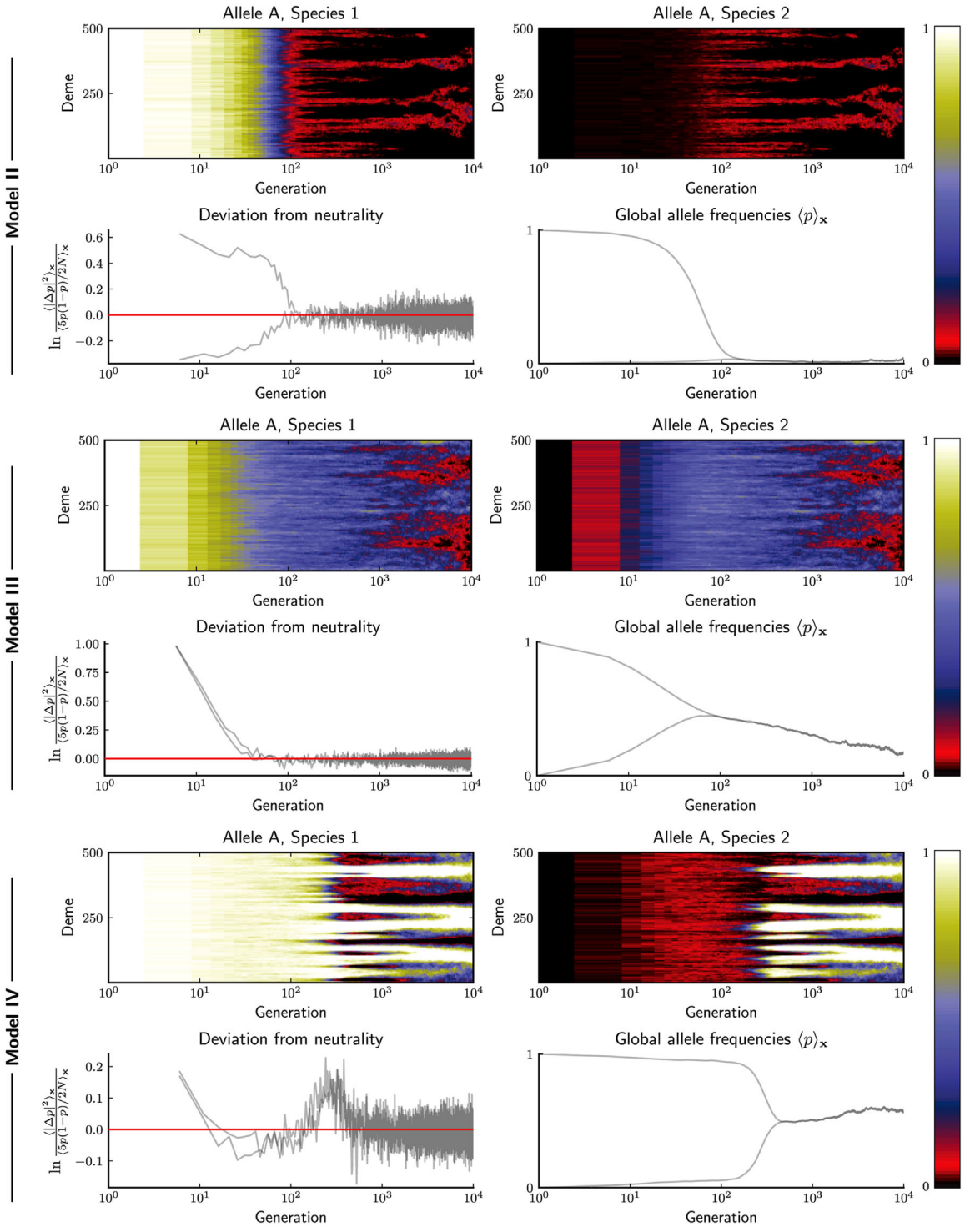

Figure 5. SLiM simulations demonstrate how redundant incompatibilities collapse into quasi-neutral polymorphisms in broadly sympatric species. Heatmaps are frequencies of allele A across space and time taken every five generations (i.e., $\Delta t = 5$ in Equation 13). For all three models, hybridization probability is $m = 0.05$, and nearest-neighbor migration probability is $m' = 0.1$. Population size in each deme fluctuates around 269. In deviation from neutrality, red lines are expectations under neutral genetic drift. See Figure S4 for the behaviors of all loci.

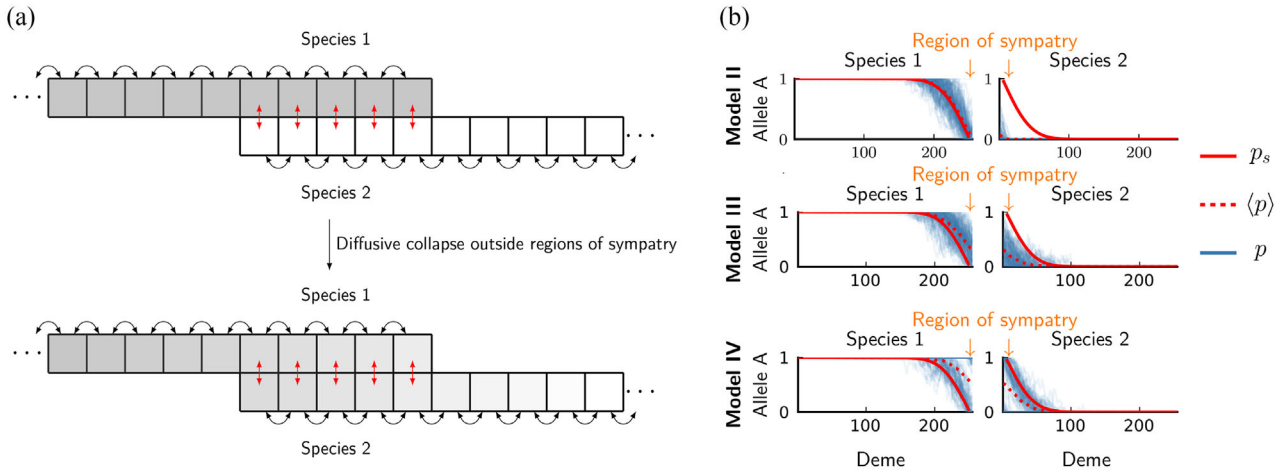

Figure 6. **The collapse of redundant incompatibilities are limited by dispersal outside regions of sympatry. (A) Schematics of the collapse outside regions of sympatry. Gray color represents frequencies of a hypothetical incompatibility allele. These frequencies form a spatial continuum determined by the dispersal process. Since frequencies are smooth, no discrete states are distinguishable (compare to nonredundant incompatibilities in Figure 4a) (B) SLiM simulations of Models II, III, and IV under a population structure where the zone of sympatric overlap has five consecutive demes, and there are 250 demes forming a one-dimensional array outside the region of sympatry in each species. Each model was run for 100 realizations, and allele frequencies (p, blue curves) were taken at generation $t = 9901$. The region of sympatry has five consecutive demes, and there are 250 demes forming a one-dimensional array outside this region in each species. Hybridization probability $m = 0.05$, and nearest-neighbor migration probability $m' = 0.1$. Solid red lines are the probability p_s calculated using Equation (14) with $\sigma^2 = m'$. Dashed red lines are the ensemble-average $\langle p \rangle$ of allele frequencies p. See Figure S5 for the behaviors of all loci.**

about 10^2 generations in Model IV of Figure 5 and Figure S4). Each bubble corresponds to a locally fixed functional pathway. In the redundant uni-stable system (Model III), pseudogenized alleles are seldom fixed locally, and their average frequencies decrease over time. However, these pseudogenes are sheltered by locally fixed functional homologs so that individuals carrying exclusively pseudogenes are rare (see banded patterns after about 10^3 generations in Model III of Figure 5 and Figure S4). The same banded pattern also occurs for Model II (see alternating black and red in Model II after about 10^2 generations in Figure 5 and Figure S4). The fact that every local deme has at least one functional component promotes quasi-neutral polymorphism in other components of the system.

The collapse of redundant incompatibilities is limited by spatial dispersal outside regions of sympatry

We have now shown that redundant incompatibilities collapse into global quasi-neutrality between broadly sympatric species. A similar quasi-neutrality argument can be applied beyond regions of sympatry because incompatibility alleles are purged indirectly there: the collapse is driven by dispersal of compatible genotypes, which gradually replace parental genotypes throughout space Figure 6a.

Consequently, the collapse rate of incompatibility alleles outside the region of sympatry is simply bounded above by the accumulation rate of compatible genotypes via dispersal. Formally, suppose an incompatibility allele is initially unique to one species and is fixed. At time t after hybridization, let $\langle p \rangle(\mathbf{x}, t)$ be the expected frequency of the incompatibility allele in that species at location $\mathbf{x}$ in the geographic range D. Let $p_s(\mathbf{x}, t)$ be the probability that a randomly sampled individual, at location $\mathbf{x} \in \mathsf{D}$ and at time t, has never hit the region of sympatry throughout its genealogical history. Then we conjecture that:

$$\langle p \rangle(\mathbf{x}, t) \geq p_s(\mathbf{x}, t), \tag{14}$$

We assume a simple model to demonstrate Equation (14). Let D be interval $[0, L]$ in one species, and suppose hybridization only affects the left boundary $x = 0$. Further, assume population density is uniform and dispersal is driven by the standard diffusion with variance σ^2 per unit time ($\sigma^2 = m'$ in discrete populations with nearest-neighbor migration). Then:

$$p_s(x, t) = \sum_{n=1}^{\infty} \frac{4/\pi}{2n-1} \sin \frac{(2n-1)\pi x}{2L} e^{-\frac{(2n-1)^2 \pi^2 \sigma^2}{8L^2} t} \tag{15}$$

Using this formula, at the right-most boundary $x = L$, the time it takes on average to decrease allele frequencies to p is lower-bounded by

$$\frac{8}{\pi^2} \left(\frac{L}{\sigma} \right)^2 \ln \frac{4}{\pi p} \tag{16}$$

The term L^2/σ^2 signifies a general pattern in one-dimensional dispersal by standard diffusion, that the rate of collapse away

from the region of sympatry slows down at least quadratically with the habitat size, which is a consequence from the scaling relationship in standard diffusion (displacement $\sim t^{\frac{1}{2}}$). By this logic, dispersal characteristics will profoundly affect the collapse rate away from the region of sympatry. If long-range dispersal occurs (i.e., a heavy-tail dispersal kernel), the collapse will be much quicker than that under standard diffusion, because incompatibility alleles are brought into the region of sympatry more quickly to be selected out. Conversely, lineages with restricted dispersal will collapse much more slowly away from the region of sympatry.

The bound p_s is compared with simulated results for Models II, III, and IV in a one-dimensional habitat with a narrow region of sympatry (Figure 6b). Despite the complexity of simulated systems, p_s curves robustly predict the "envelope" of allele frequency clines, and the ensemble-average $\langle p \rangle$ of all allele frequency clines are always bounded by p_s. Thus, the conjecture holds under a simple habitat model, and the rate of collapse away from the region of sympatry can be largely predicted from spatial dispersal alone.

Discussion

Multilocus epistasis can be formidably complex. Here, we derive general predictions about the collapse of interacting incompatibilities in nature by focusing on the level of their genetic redundancy. The key to understanding this classification based on genetic redundancy is that many empirical incompatibility systems appear to have evolved via duplications of genes or pathways, followed by divergent "developmental system drift" (True & Haag, 2001) in separate populations. We show that different levels of genetic redundancy will lead to qualitatively different behaviors of incompatibilities under hybridization. Provided hybridization is rare, redundant incompatibilities collapse continuously over time, often associated with various quasi-neutral behaviors in space and time. Nonredundant incompatibilities, on the other hand, exhibit bistable (or multistable) dynamics, which lead to abrupt collapse (fitness peak shift, and traveling waves or tension zones in a spatial system), and will be more sensitive to stochastic events. In particular, while redundancy renders the system unstable to gene flow, it also serves as a buffer that prevents the complete extinction of incompatibility alleles, and the likelihood of widespread polymorphism could be high with redundant incompatibilities in spatially extended populations.

THE IMPACT OF ECOLOGICAL SELECTION

Alleles involved in incompatibilities might often be adaptive (Coyne & Orr, 2004). However, two types of adaptive alleles should be recognized: (1) Incompatibility alleles such as suppressors of meiotic drive that arise mainly to attenuate perturbations to an organism's intrinsic physiology; (2) Incompatibility alleles under ecological selection (local adaptation). We have ignored the second type in the analysis of collapse dynamics, but the overall effect of ecological selection is not difficult to predict—incompatibility alleles strongly favored by a particular local environment will be hard to purge. To this end, a number of models have been analyzed to find the critical point where local adaption and incompatibility jointly prevent the collapse of species barriers in a single population receiving migrants (Bank et al., 2012; Blanckaert & Hermisson, 2018; Blanckaert et al., 2020). In our theory, since hybridization is assumed to be a weak force, local adaptation could alter collapse dynamics in the following ways:

- If incompatibilities are redundant, quasi-neutrality might not hold because local adaptation could cause breaks in the plateau of the fitness landscape (Figure 2b).
- If incompatibilities are non-redundant, local adaptation could alter the height of each fitness peak, thus increasing the asymmetry of incompatibilities in each local population.

With additional simulations, we show that a weak ecological advantage ($s \sim 0.01$) does not qualitatively alter the abrupt collapse of nonredundant incompatibilities (Figure S6). The quantitative difference is that, with larger s, the waiting time to a local collapse increases, and the number of demes experiencing local collapse becomes smaller, but traveling waves still propagate once they are initiated. For redundant incompatibilities between broadly sympatric species, weak ecological selection has a negligible impact on the short term collapse (Figure S7– S9), because the fitness cost of incompatibility overshadows the weak levels of local adaptation. In the long term, after incompatibility collapses, spatial clusters of genotypes gradually emerge as a result of ecological selection, and the entire system still remains close to quasi-neutral polymorphism. Nonetheless, ecological selection has a more complex influence on incompatibilities across a narrow region of sympatry (Figure S10). Because the fitness cost of incompatibility is not "felt" outside the region of overlap, so that adaptive incompatibility alleles could be locally fixed, which can facilitate the formation of stable clines (Model II and IV in Figure S10).

An open question is how often are incompatibility alleles under ecological selection? Systems involving pseudogenes perhaps have little connection to ecological performance. Incompatibilities attributed to selfish genetic elements are also not connected to ecology in obvious ways. However, immune-response alleles causing hybrid necrosis might be strongly favored when specific pathogens are abundant, and highly polygenic incompatibilities could be involved in ecological adaptation solely because of the large number of genes. Species maintained solely by ecological selection are vulnerable to hybridization following environmental

change (Taylor et al., 2006; Vonlanthen et al., 2012), and therefore perhaps do not persist for long.

REVISITING INCOMPATIBILITY ACCUMULATION

Orr's influential model of incompatibility accumulation predicts a "snowball" effect, that incompatibilities appear faster than linearly with time (Orr & Turelli, 2001). This conjecture assumes that multilocus interactions evolve locus-by-locus independently in isolated populations, and that epistatic incompatibilities become more likely as the number of substitutions increases. In real genetic/biochemical networks, components of statistical epistasis are not independent of one another, and gene network topology might strongly influence fitness landscapes, and hence the rate of incompatibility accumulation (Guerrero et al., 2017).

For a tightly connected, interdependent genetic network, genetic redundancy is low, and the system will be sensitive to disruptions. Hence, establishing nonredundant incompatibilities might appear difficult, since two parental genotypes are on different fitness peaks separated by a large fitness trough. Crossing the fitness trough stochastically via large-effect mutations will likely fail in populations of reasonable size (Coyne et al., 1997). However, this difficulty can be circumvented via developmental system drift (Schiffman & Ralph, 2022; True & Haag, 2001), where interacting loci accumulate small, compensatory changes (e.g., gradually altering sequences of transcription factors and their binding sites) under stabilizing selection. These processes (which do not necessarily require genetic drift in the ordinary sense) can eventually lead to large two-locus fitness troughs between divergent populations. Another possible route will involve duplication of some components of the network (redundancy ↑) so that redundant components can diverge freely, followed by the loss of duplicated components after divergence (redundancy ↓) (see fig. 1 in Haag and Molla 2005). This suggests that changes in network topology that are closely connected to levels of genetic redundancy could play a role in accumulation dynamics as well as in collapse dynamics.

CONCLUSION: TWISTED ROUTES TO SPECIATION?

The species concept is open to many interpretations, but a bifurcating tree describing species relationships remains a widely accepted picture since Charles Darwin, even in this genomic age where reticulation has been found to be a common feature (Mallet et al., 2016). The pictorial simplicity of this macroscopic process does not preclude a wide variety of microscopic mechanisms operating near the bifurcation point. In this work, we show that intrinsically incompatible species pairs can still collapse in various ways due to hybridization, leading to speciation reversal if genetic incompatibility is adopted as a species criterion. In practice, genetically compatible populations may still mate assortatively, live under different ecological conditions, and show high linkage disequilibrium, so that they can retain their species status even after intrinsic incompatibilities have collapsed. In most cases, unambiguously splitting lineages into distinct species will be achieved only after multiple speciation traits coincide (e.g., strongly incompatible lineages with assortative mating and ecological differentiation). Therefore, as a particular variable in speciation, genetic incompatibility should be treated as an accompanying phenomenon that is mostly positively correlated with the speciation progress, but could also reverse direction between incipient lineages. Detecting such process in empirical systems and assessing its prevalence will help us achieve a more objective understanding of what happens near the bifurcation point that we call speciation.

AUTHOR CONTRIBUTIONS

T.X. and J.M. conceived the project; T.X. constructed and analyzed the theory; T.X. and J.M. wrote the manuscript.

ACKNOWLEDGMENTS

T.X. was funded by a studentship from the Department of Organismic and Evolutionary Biology; the NSF-Simons Center for Mathematical and Statistical Analysis of Biology (award number #1764269); and the Quantitative Biology Initiative, Harvard University. We thank John Wakeley for detailed comments on earlier drafts of this paper. We thank Robin Hopkins and Naomi Pierce for general feedback, Nathaniel Edelman and Neil Rosser for discussion of the incompatiblities between *Heliconius* butterflies which inspired the current work.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

DATA ARCHIVING

Simulation codes and raw data are archived in Dryad: https://doi.org/10.5061/dryad.mgqnk992f

REFERENCES

AkhavanAghdam, Z., Sinha, J., Tabbaa, O.P. & Hao, N. (2016) Dynamic control of gene regulatory logic by seemingly redundant transcription factors. *eLife*, 5, e18458.

Atlan, A., Capillon, C., Derome, N., Couvet, D. & Montchamp-Moreau, C. (2003) The evolution of autosomal suppressors of sex-ratio drive in *Drosophila simulans*. *Genetica*, 117(1), 47–58.

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

Barolo, S. (2012) Shadow enhancers: frequently asked questions about distributed *cis*-regulatory information and enhancer redundancy. *Bioessays*, 34(2), 135–141.

Barriere, A., Gordon, K.L. & Ruvinsky, I. (2012) Coevolution within and between regulatory loci can preserve promoter function despite evolutionary rate acceleration. *PLoS Genetics*, 8(9)

Barton, N. (1979) The dynamics of hybrid zones. *Heredity*, 43, 341–359.

Barton, N. & Bengtsson, B.O. (1986) The barrier to genetic exchange between hybridising populations. *Heredity*, 57(3), 357–376.

Barton, N.H., Etheridge, A.M. & Véber, A. (2013) Modelling evolution in a spatial continuum. *Journal of Statistical Mechanics: Theory and Experiment*, 2013(01), P01002.

Barton, N.H. & Turelli, M. (2011) Spatial waves of advance with bistable dynamics: cytoplasmic and genetic analogues of Allee effects. *The American Naturalist*, 178(3), E48–E75.

Bateson, W. (1909) Heredity and variation in modern lights. In A. Seward (Ed.), Darwin and Modern Science (pp. 85–101). Cambridge University Press.

Bazykin, A. (1973) Population genetical analysis of disruptive and stabilizing selection. ii. adjacent populations with continuous area. *Genetika, Moskva*, 9, 156–166.

Blanckaert, A., Bank, C. & Hermisson, J. (2020) The limits to parapatric speciation 3: Evolution of strong reproductive isolation in presence of gene flow despite limited ecological differentiation. *Philosophical Transactions of the Royal Society B*, 375(1806), 20190532.

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

Case, A.L., Finseth, F.R., Barr, C.M. & Fishman, L. (2016) Selfish evolution of cytonuclear hybrid incompatibility in *Mimulus*. *Proceedings of the Royal Society B: Biological Sciences*, 283(1838), 20161493.

Chae, E., Bomblies, K., Kim, S.T., Karelina, D., Zaidem, M., Ossowski, S., Martín-Pizarro, C., Laitinen, R.A., Rowan, B.A., Tenenboim, H., et al. (2014) Species-wide genetic incompatibility analysis identifies immune genes as hot spots of deleterious epistasis. *Cell*, 159(6), 1341–1351.

Chou, J.Y. & Leu, J.Y. (2010) Speciation through cytonuclear incompatibility: insights from yeast and implications for higher eukaryotes. *Bioessays*, 32(5), 401–411.

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

Cotton, J.A. & Page, R.D. (2005) Rates and patterns of gene duplication and loss in the human genome. *Proceedings of the Royal Society B: Biological Sciences*, 272(1560), 277–283.

Coughlan, J.M. & Matute, D.R. (2020) The importance of intrinsic postzygotic barriers throughout the speciation process. *Philosophical Transactions of the Royal Society B*, 375(1806), 20190533.

Courret, C., Gérard, P.R., Ogereau, D., Falque, M., Moreau, L. & Montchamp-Moreau, C. (2019) X-chromosome meiotic drive in *Drosophila simulans*: a QTL approach reveals the complex polygenic determinism of Paris drive suppression. *Heredity*, 122(6), 906–915.

Coyne, J. & Orr, H. (2004) Speciation. Sunderland, Mass: Sinauer Associates.

Coyne, J.A., Barton, N.H. & Turelli, M. (1997) Perspective: a critique of sewall wright's shifting balance theory of evolution. *Evolution*, 51(3), 643–671.

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

Dagilis, A.J., Kirkpatrick, M. & Bolnick, D.I. (2019) The evolution of hybrid fitness during speciation. *PLoS Genetics*, 15(5), e1008125.

Dobzhansky, T. (1937) Genetics and the origin of species. New York: Columbia University Press.

Gavrilets, S. (1997) Evolution and speciation on holey adaptive landscapes. *Trends in Ecology & Evolution*, 12(8), 307–312.

Geldhauser, C. & Kuehn, C. (2020) Travelling waves for discrete stochastic bistable equations. *arXiv*.

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

Haag, E.S. (2007) Compensatory vs. pseudocompensatory evolution in molecular and developmental interactions. *Genetica*, 129(1), 45–55.

Haag, E.S. & Molla, M.N. (2005) Compensatory evolution of interacting gene products through multifunctional intermediates. *Evolution*, 59(8), 1620–1632.

Haller, B.C. & Messer, P.W. (2019) SLiM 3: Forward genetic simulations beyond the Wright–Fisher model. *Molecular Biology and Evolution*, 36(3), 632–637.

Hatfield, T., Barton, N. & Searle, J.B. (1992) A model of a hybrid zone between two chromosomal races of the common shrew (*Sorex araneus*). *Evolution*, 46(4), 1129–1145.

Hu, Z., Killion, P.J. & Iyer, V.R. (2007) Genetic reconstruction of a functional transcriptional regulatory network. *Nature Genetics*, 39(5), 683–687.

Jiao, W.B., Patel, V., Klasen, J., Liu, F., Pecinkova, P., Ferrand, M., Gy, I., Camilleri, C., Effgen, S., Koornneef, M., et al. (2021) The evolutionary dynamics of genetic incompatibilities introduced by duplicated genes in *Arabidopsis thaliana*. *Molecular Biology and Evolution*, 38(4), 1225–1240.

Josefsson, C., Dilkes, B. & Comai, L. (2006) Parent-dependent loss of gene silencing during interspecies hybridization. *Current Biology*, 16(13), 1322–1328.

Larson, E.L., Vanderpool, D., Sarver, B.A., Callahan, C., Keeble, S., Provencio, L.L., Kessler, M.D., Stewart, V., Nordquist, E., Dean, M.D., et al. (2018) The evolution of polymorphic hybrid incompatibilities in house mice. *Genetics*, 209(3), 845–859.

Liénard, M.A., Araripe, L.O. & Hartl, D.L. (2016) Neighboring genes for DNA-binding proteins rescue male sterility in *Drosophila* hybrids. *Proceedings of the National Academy of Sciences*, 113(29), E4200–E4207.

Lindtke, D. & Buerkle, C.A. (2015) The genetic architecture of hybrid incompatibilities and their effect on barriers to introgression in secondary contact. *Evolution*, 69(8), 1987–2004.

Lynch, M. & Force, A.G. (2000) The origin of interspecific genomic incompatibility via gene duplication. *The American Naturalist*, 156(6), 590–605.

Mack, K.L. & Nachman, M.W. (2017) Gene regulation and speciation. *Trends in Genetics*, 33(1), 68–80.

Mallet, J. (2005) Hybridization as an invasion of the genome. *Trends in Ecology & Evolution*, 20(5), 229–237.

Mallet, J., Besansky, N. & Hahn, M.W. (2016) How reticulated are species? *BioEssays*, 38(2), 140–149.

Meiklejohn, C.D., Landeen, E.L., Gordon, K.E., Rzatkiewicz, T., Kingan, S.B., Geneva, A.J., Vedanayagam, J.P., Muirhead, C.A., Garrigan, D., Stern, D.L., et al. (2018) Gene flow mediates the role of sex chromosome meiotic drive during complex speciation. *eLife*, 7, e35468.

Moore, R.C. & Purugganan, M.D. (2005) The evolutionary dynamics of plant duplicate genes. *Current Opinion in Plant Biology*, 8(2), 122–128.

Morán, T. & Fontdevila, A. (2014) Genome-wide dissection of hybrid sterility in *Drosophila* confirms a polygenic threshold architecture. *Journal of Heredity*, 105(3), 381–396.

Morimoto, K., Numata, K., Daitoku, Y., Hamada, Y., Kobayashi, K., Kato, K., Suzuki, H., Ayabe, S., Yoshiki, A., Takahashi, S., et al. (2020) Reverse genetics reveals single gene of every candidate on *Hybrid sterility, X Chromosome QTL 2* (*Hstx2*) are dispensable for spermatogenesis. *Scientific Reports*, 10(1), 1–9.

Muller, H. (1942) Isolating mechanisms, evolution, and temperature. In Biol. Symp., volume 6, (pp. 71–125).

Nowak, M.A., Boerlijst, M.C., Cooke, J. & Smith, J.M. (1997) Evolution of genetic redundancy. *Nature*, 388(6638), 167–171.

Orr, H.A. & Turelli, M. (2001) The evolution of postzygotic isolation : accumulating Dobzhansky-Muller incompatibilities. *Evolution*, 55(6), 1085–1094.

Phadnis, N. & Orr, H.A. (2009) A single gene causes both male sterility and segregation distortion in *Drosophila* hybrids. *Science*, 323(5912), 376–379.

Poelwijk, F.J., Tănase-Nicola, S., Kiviet, D.J. & Tans, S.J. (2011) Reciprocal sign epistasis is a necessary condition for multi-peaked fitness landscapes. *Journal of Theoretical Biology*, 272(1), 141–144.

Porter, A.H. & Johnson, N.A. (2002) Speciation despite gene flow when developmental pathways evolve. *Evolution*, 56(11), 2103–2111.

Prager, E.M. & Wilson, A.C. (1975) Slow evolutionary loss of the potential for interspecific hybridization in birds: a manifestation of slow regulatory evolution. *Proceedings of the National Academy of Sciences*, 72(1), 200–204.

Presgraves, D.C. & Meiklejohn, C.D. (2021) Hybrid sterility, genetic conflict and complex speciation: lessons from the *Drosophila simulans* clade species. *Frontiers in Genetics*, 12.

Rosser, N., Edelman, N.B., Queste, L.M., Nelson, M., Seixas, F., Dasmahapatra, K.K. & Mallet, J. (2021) Complex basis of hybrid female sterility and Haldane's rule in *Heliconius* butterflies: Z-linkage and epistasis. *Molecular Ecology*.

Rosser, N., Queste, L.M., Cama, B., Edelman, N.B., Mann, F., Mori Pezo, R., Morris, J., Segami, C., Velado, P., Schulz, S., et al. (2019) Geographic contrasts between pre-and postzygotic barriers are consistent with reinforcement in *Heliconius* butterflies. *Evolution*, 73(9), 1821–1838.

Schiffman, J.S. & Ralph, P.L. (2022) System drift and speciation. *Evolution*, 76(2), 236–251.

Schlosser, G. & Wagner, G.P. (2004) Modularity in development and evolution. University of Chicago Press.

Seidel, H.S., Rockman, M.V. & Kruglyak, L. (2008) Widespread genetic incompatibility in *C. elegans* maintained by balancing selection. *Science*, 319(5863), 589–594.

Taylor, E., Boughman, J., Groenenboom, M., Sniatynski, M., Schluter, D. & Gow, J. (2006) Speciation in reverse: morphological and genetic evidence of the collapse of a three-spined stickleback (*Gasterosteus aculeatus*) species pair. *Molecular Ecology*, 15(2), 343–355.

True, J. & Haag, E. (2001) Developmental system drift and flexibility in evolutionary trajectories. *Evolution and Development*, 3, 109–119.

Vonlanthen, P., Bittner, D., Hudson, A.G., Young, K.A., Müller, R., Lundsgaard-Hansen, B., Roy, D., Di Piazza, S., Largiader, C.R. & Seehausen, O. (2012) Eutrophication causes speciation reversal in whitefish adaptive radiations. *Nature*, 482(7385), 357–362.

Wright, S. (1932) The roles of mutation, inbreeding, crossbreeding, and selection in evolution. *Proceedings of the VI International Congress of Genetics, The Hague*, 1, 356–366.

Wright, S. (1943) Isolation by distance. *Genetics*, 28(2), 114.

Xiong, T. & Mallet, J. (2022) On the impermanence of species: the collapse of genetic incompatibilities in hybridizing populations. Dryad, Dataset.

Zuellig, M.P. & Sweigart, A.L. (2018a) Gene duplicates cause hybrid lethality between sympatric species of *Mimulus*. *PLoS Genetics*, 14(4), e1007130.

Zuellig, M.P. & Sweigart, A.L. (2018b) A two-locus hybrid incompatibility is widespread, polymorphic, and active in natural populations of *Mimulus*. *Evolution*, 72(11), 2394–2405.

Associate Editor: Kavita Jain
Handling Editor: Tracey Chapman

Supporting Information

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

Figure S1: Comparison of the collapse time in Models I & II between analytic approximations and simulated birth-death prcesses
Figure S2: One-dimensional asymptotic equations like Equations (S9) and (S27) are good approximations to the full haplotype dynamics under strong incompatibilities and weak hybridization in Models I & II and their variants.
Figure S3: Travelling waves: behaviors of all loci in Model I between two broadly-sympatric species.
Figure S4: Quasi-neutral polymorphism: behaviors of all loci in Model II, III, and IV, between two broadlysympatric species.
Figure S5: Quasi-neutral polymorphism: behaviors of all loci in Model II, III, IV, between two parapatric species with a narrow region of sympatry.
Figure S6: Ecological selection: the collapse dynamics in Model I between broadly sympatric species.
Figure S7: Ecological selection: the collapse dynamics in Model II between broadly sympatric species.
Figure S8: Ecological selection: the collapse dynamics in Model III between broadly sympatric species
Figure S9: Ecological selection: the collapse dynamics in Model IV between broadly sympatric species.
Figure S10: Ecological selection: behaviors of all loci under ecological selection in Model II, III, IV, between two species with a narrow region of sympatry (sympatric in only 5 demes, with additional 250 demes flanking on each side.